Acc.No. 5119

1. spectroscopy 2. Chromatography
3. electroencephalography
4. ultracentrifuge 5. Xray
6. fluorescent antibody
7. phonocardiography (10) microradiography
8. vectorcardiography
9. electromyography

TLM

TOOLS OF BIOLOGICAL RESEARCH

SECOND SERIES

TOOLS OF BIOLOGICAL RESEARCH

SECOND SERIES

Edited by

HEDLEY J. B. ATKINS

D.M., M.Ch., F.R.C.S. Eng., Hon. F.A.C.S.

Director of the Department of Surgery, Guy's Hospital
Dean of the Institute of Basic Medical Sciences

With an Introduction by

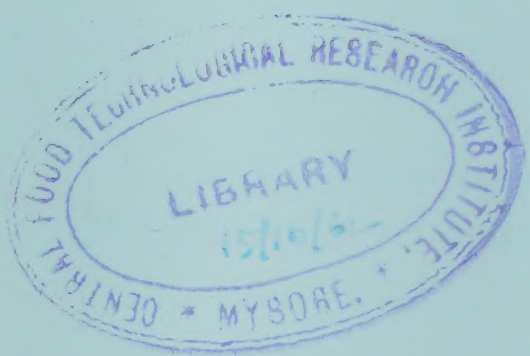
LORD COHEN OF BIRKENHEAD

BLACKWELL
SCIENTIFIC PUBLICATIONS
OXFORD

Published simultaneously in the United States of America by Charles C Thomas, Publisher, 301–327 East Lawrence Avenue, Springfield, Illinois.

Published simultaneously in Canada by the Ryerson Press, Queen Street West, Toronto 2.

FIRST PUBLISHED AUGUST 1960



PRINTED IN ENGLAND BY ADLARD & SON LTD
BARTHOLOMEW PRESS, DORKING
AND BOUND BY THE KEMP HALL BINDERY, OXFORD

CONTENTS

EDITOR'S PREFACE

The first Symposium on 'Tools of Biological Research' was held in 1958 at Guy's Hospital primarily for the benefit of members of the Surgical Research Society together with certain other research workers who were invited to attend, and a collection of the papers given there has already been published. This conference was so successful, and the number of 'Tools' which had to be left out of the symposium so numerous, that a further conference was held in 1959 to cover some of the important ones. Again the subjects chosen were such as would be of interest to surgeons engaged in research, and the manner in which these subjects were treated by those who gave papers was dictated by these interests. Nevertheless, the subjects are of wide applicability and it is hoped that the present publication which is a collection of the papers given on this second occasion will also be of value to research workers in other fields of applied biology.

The second symposium was introduced by Lord Cohen of Birkenhead whose address serves as a fitting introduction to the series. It was a great privilege to be able to welcome him and his attendance was an attestation of a lifelong interest in the scientific basis of medicine as well as an encouragement to us to think that the project was worth while.

Since the publication of the first volume there has been a substantial rise in salaries within the printing industry and a general increase in costs; it was found impossible therefore without help to publish the present volume at the same price as the first. The publishers from whom I have, as usual, received every consideration and kindness, were distressed that the number of blocks, together with coloured illustrations made it impossible for them to bring the price of the book within the range of those for whom it was intended. An appeal was made to the Welcome Trust who made a most generous grant towards the cost and have enabled this, the more expensive production, to be sold at the same price as the former publication.

Certain illustrations have already appeared elsewhere and acknowledgements have been made in the appropriate places in the text. If, unwittingly, some acknowledgement has been omitted, I find from experience that this is almost certainly due to the fact that the author has forgotten that he has previously published a certain illustration elsewhere. It is perhaps

an old friend that he uses as a slide to illustrate his lectures and has come through long usage to be regarded as his own property. As editor I must apologize for such deficiencies in advance and give my assurances that no disrespect was intended.

Finally I wish to thank the contributors who, without reward other than that of earning our gratitude, have been at such pains to explain in as simple language as possible the intricacies of their craft.

Guy's Hospital. Hedley Atkins.

1960

INTRODUCTION

LORD COHEN OF BIRKENHEAD

When the history of medicine during the past hundred years comes to be written, the dominant theme will I suggest be the disappearance of the ontological concept of disease. Of this Sydenham was the supreme interpreter. In his own words diseases were 'to be reduced to certain and determinate kinds, with the same exactness as we see it done by botanic writers in their treatises of plants', a taxonomic approach which was to reach its climax in Linnaeus's dictum a century later—'*Species tot sunt diversae quot diversae formae ab initio sunt creatae.*'

This concept was appropriate for a time when the physician was a naturalist; when his only tools of biological research were his five senses. Then he discerned and described, sometimes with outlines blurred, the uniformly recurring patterns of gout, of measles, of smallpox, of chorea and like 'diseases', and regarded them as entities, and here I again quote Sydenham, to be recognized by 'certain distinguishing signs which Nature has particularly affixed to every species'. Thus the ill patient was himself *plus* a disease. In a limited sense this view appeared to receive support when the bacterial causes were discovered of many of the diseases which Sydenham and his predecessors recognized. But despite this temporary fillip in the latter part of the nineteenth century, this concept of disease 'entities' has virtually passed and has given place to the current view that the more rational and fruitful approach to the problem of disease is to regard it as a disturbance of the normal structure and function of the body brought about by manifold agents, both extrinsic and intrinsic (including genetic propensities) which act not in isolation but in concert, though in varying degrees of partnership.

This retreat from Sydenham demanded an act of faith, which arose from the belief that the living organism is in many (and the monist would affirm in all) respects a physico-chemical mechanism, whose workings can and will be elucidated by the methods and instruments of physics and chemistry. This is no novel thesis although it has taken nearly four centuries for it to be generally accepted. Its origins antedate Sydenham

by nearly a century. They can be traced to Pisa, where, in the second half of the sixteenth century, Galileo observed and then verified experimentally the isochronism of the swing of the pendulum and applied it to the timing of the human pulse, and proceeding to show 'order amid change', to use Whitehead's phrase, he demonstrated by measurement the laws of falling bodies, of equilibrium and of motion on an inclined plane, and of the parabolic motion of a projectile. It was when Galileo was at Padua where he taught from 1591 to 1610, that he inspired Sanctorius to devise tools which might reveal similar order in man and it is to him that we owe, for example, the earliest 'clinical' thermometer, and an apparatus for comparing pulse rates; and he built a sufficiently accurate weighing machine for him to discern the 'insensible perspiration' of the body which exceeds all other bodily excretions combined—the first recorded metabolic balance test.

In the seventeenth century the transformation of alchemy into the science of chemistry, which was initiated by Robert Boyle and John Mayow, led to the investigation by then available methods, notably by Franciscus Sylvius and his school, of the chemistry of the saliva and digestive juices.

This tide of materialism was stemmed for a time at the turn of the eighteenth century by George Ernst Stahl (1660–1734), whose theory of 'vitalism' insisted that the phenomena exhibited by living organisms are not governed by physical and chemical laws but by the 'sensitive soul'—the 'psyche' of Aristotle, on a plane far above physics and chemistry. But in its original form it could not withstand the blow aimed at it by Wöhler's synthesis of urea in 1828—a blow to the vitalists more widely recognized than was Harvey's of two centuries earlier when, before his experimental approach to the problem he tackled in *De Motu Cordis*, he was, and here I quote, 'almost tempted to think, with Frascatorius, that the motion of the heart was only to be comprehended by God.'

In Italy also electrophysiology was born towards the end of the eighteenth century by the chance observation of Galvani and the consequential experiments of Volta.

From these small beginnings over three and a half centuries ago have sprung especially during the past half century the prodigious array of instruments and techniques of physics and chemistry often of bewildering complexity and incredible potentiality, which are being applied to the study of biological phenomena and some of which are to be discussed in this symposium.

From Volta's experiments to modern electromyography, electro-

cardiography and electroencephalography might well appear a long span in space and time, but the milestones are direct, continuous, consequential and easily discerned, and we can be confident that the recent rapid and stupendous progress in electronic equipment and techniques means not only more accurate recording and analysis of electrical changes in tissues and fluids, but because of the size and simplicity of these newer instruments a wider field of research will assuredly be opened. The rate of progress in this field is acceleration unrestrained. It is only sixty-four years since Röntgen announced the discovery of X-rays; yet today X-ray microscopy has advanced to a stage whence we look forward with confidence to the day in the near future when by differential absorption measurements, the nature and concentration of elements in any specimen will be given, and if the constituents are crystalline their identification will be completed by microdiffraction.

In this symposium the methods, scope and potentialities of some of these newer weapons in the battle against nature's secrets will be discussed, and it may be that you will become aware that they have a contribution to make to the problems on which you are currently engaged. You will certainly be strengthened in the view that virtually every advance in physico-chemical theory and techniques has a part to play in biological research.

But, in research, instruments of themselves are not enough; there must also be a clear objective, perspicacious planning, and concentrated effort. And there are many lessons still to be learnt. Perhaps the most important is that the complexity of modern techniques demands a team of experts in many disciplines; the biologist cannot work in isolation. Amongst others, Norbert Wiener's work on cybernetics and communication theory convincingly exemplifies the view that if we are to use modern instruments to the best advantage we, who are essentially human biologists, must work with those best able to give us help in interpreting our results; otherwise, for example, the ultracentrifuge might yield for us empirically useful but rationally worthless curves. Again, the instrument and the problem must be well suited to one another; the various forms of electrophoresis and chromatography and many of the other techniques to be discussed have each their optimal applications. As A. D. Ritchie observed —'The right way to catch sharks is not the right way to catch herrings.' Those who ignore his dictum are prone to fish up old shoes. And lastly, it is well to recall that all knowledge is not yet ours. It may be that no technique or instrument has yet been devised for the solution of your special problem. It may be that some secrets, for example, that of the

cancer cell, will be locked in Nature's keeping until some key, as yet undreamt of, falls into the hands of one who has the wit to recognize it.

Wordsworth discerned nearly a century ago that

> 'Nature hides
> Her secrets less and less; Man now presides
> In power, where once he trembled in his weakness;
> Science advances with gigantic strides.'

This symposium is designed to make more widely known the tools by which these gigantic strides are today being taken in the field of the biological sciences, and it is for me a most acceptable privilege to launch it on what I am sure will inaugurate many rewarding journeys.

ELECTRON SPIN RESONANCE SPECTROSCOPY

S. J. Wyard

The first experiments in Electron Spin Resonance Spectroscopy were carried out independently and almost simultaneously in Russia and in the United States of America, at the end of the last war; it is pleasant to note that much of the subsequent development has been done in this country. At first the spectroscopes were used almost exclusively by physicists; later they were used for chemical problems; and in the last few years biologists and biochemists have been applying this new tool to their own fields of research.

Physical Principles

E.S.R. spectroscopy has much in common with absorption spectroscopy in the visible, ultra-violet and infra-red regions. In each case a beam of electromagnetic radiation is passed through the material under investigation, and it is found that the radiation is absorbed more at certain wave-lengths than at others. A plot of absorption against wave-length is called the absorption spectrum; and the information in the spectrum can be interpreted to give information about the atoms and molecules present in the material, how they are arranged, and what reactions, if any, are taking place. E.S.R. spectroscopy differs from these other forms of absorption spectroscopy in that the material is placed in a magnetic field while the absorption spectrum is being obtained. To see the reason for this, and the effect of the magnetic field, we must have a brief look at a fundamental part of physics.

The basis of E.S.R. spectroscopy is the spinning electron. The atom is sometimes compared to the solar system, as both the solar system and the atom have a central heavy body, sun or nucleus, round which the other lighter bodies, planets or electrons, move in their orbits. The comparison may be carried a step further than this, because the electrons, like the earth, spin on their axes and are magnetic. The combination of these two properties, rotation about the axis and magnetic moment, goes

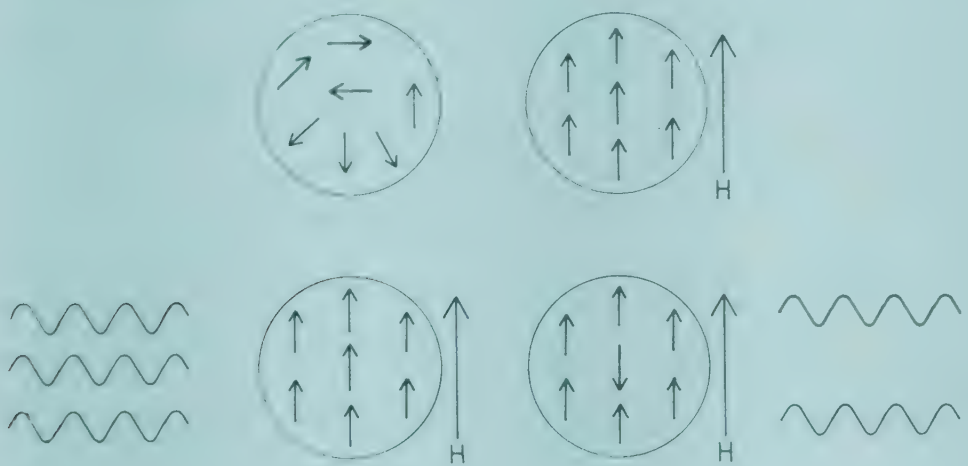
by the name of 'spin'. However, unlike any model one can imagine, the electron has the peculiar property that when placed in a magnetic field its axis can only lie in one of two directions, either parallel to the direction of the field or anti-parallel. Perhaps the model which most nearly approaches the spinning electron is the toy known as the Tippe-top, which spins either on its head or its base, according to the speed of rotation.

The electron spinning in a magnetic field has more energy with its axis anti-parallel to the field direction than when it is parallel to the field direction. If exactly the right amount of energy is added (or subtracted) to an electron spinning in a magnetic field, the electron will turn upside down. This reversal can be achieved if electromagnetic radiation of the right frequency is applied in the right direction. The correct frequency, or frequency for resonance, is directly proportional to the strength of the magnetic field. In principle any frequency could be used, with the corresponding magnetic field. In practice a high frequency is generally used, because then the spectroscope is more sensitive; and most E.S.R. spectroscopes work at radar frequencies of about 9,000 megacycles per second.

The vast majority of all electrons go in pairs, with the two spins pointing in opposite directions. A pair of electrons cannot be detected in the spectroscope, because if one electron turns round in the magnetic field, its partner does also, with no net change. Consequently only single, or unpaired electrons can be detected, and most materials give no signal when placed in the spectroscope. Unpaired electrons occur in two ways, firstly in the paramagnetic elements, e.g. iron, copper and manganese, and secondly in the chemical species known as free radicals. Most applications of E.S.R. spectroscopy to biological problems so far have dealt with this second group, the free radicals. A free radical may be defined as a molecule, or part of a molecule, in which one or more of the valence electrons (i.e. the electrons which take part in chemical binding) is unpaired. The E.S.R. spectroscope should be able to tell us whether free radicals are present in a given material, and if so, their concentration and lifetime. Sometimes, but not always, it is possible to identify the free radicals from the E.S.R. spectrum.

If a specimen containing free radicals is placed in the spectroscope and the magnetic field is varied while the frequency of the electromagnetic radiation is kept constant (this is usually more convenient than the alternative of varying the frequency while the magnetic field is kept constant) then we should expect the power absorbed by the specimen to undergo a sudden increase at the resonant conditions, because the radiation

is turning the electrons round against the magnetic field. The situation is represented in diagrammatic form by Fig. 1. A plot of power absorbed against magnetic field strength, which is the usual way of displaying the spectrum in E.S.R. spectroscopy, would look like Fig. 2(*a*). The height of the line is proportional to the number of free radicals in the sample, but the position of the line, i.e. the value of the magnetic field at which resonance occurs (which is generally expressed in terms of the 'g-value') would be the same for all free radicals. Now we have assumed so far that the unpaired electrons are affected by the magnetic field, but are otherwise free. In fact every unpaired electron is bound to a molecule, and is in-



$$E = h\nu = g\beta H$$

$$\nu \simeq 2{\cdot}8H \text{ Mc/s (H IN GAUSS)}$$

Figure 1. Diagram of E.S.R. The arrows represent unpaired electrons and the waves represent radiation. When a magnetic field H is applied the electrons tend to be lined up in the same direction. If radiation of the right frequency ν is passed through the sample, some radiation is absorbed in turning round the electrons against the magnetic field. The value of ν is given by the equations, where h and β are constants, but g varies from one material to another. The numerical value corresponds to a 'free electron'.

fluenced by the surrounding molecules and other unpaired electrons. The result of this is to broaden the absorption line and to alter the g-value (which is now the value for which the absorption is at a maximum) so that the spectrum looks like Fig. 2(*b*). The line-width (width of the curve at half height) and g-value vary from specimen to specimen, and often give useful information, but will not usually identify the free

radical. This is because many factors affect the line-width, including such extraneous ones as the temperature and viscosity of the specimen; and because the variation in the g-value is usually less than $\frac{1}{2}$ per cent.

Fortunately there is often another means of identifying free radicals from the absorption spectrum, by what is known as 'hyperfine splitting'. This arises from the fact that many nuclei are themselves magnetic. These nuclei take up certain directions with respect to the applied magnetic field, just as electrons do. If the orbit of an unpaired electron takes it close

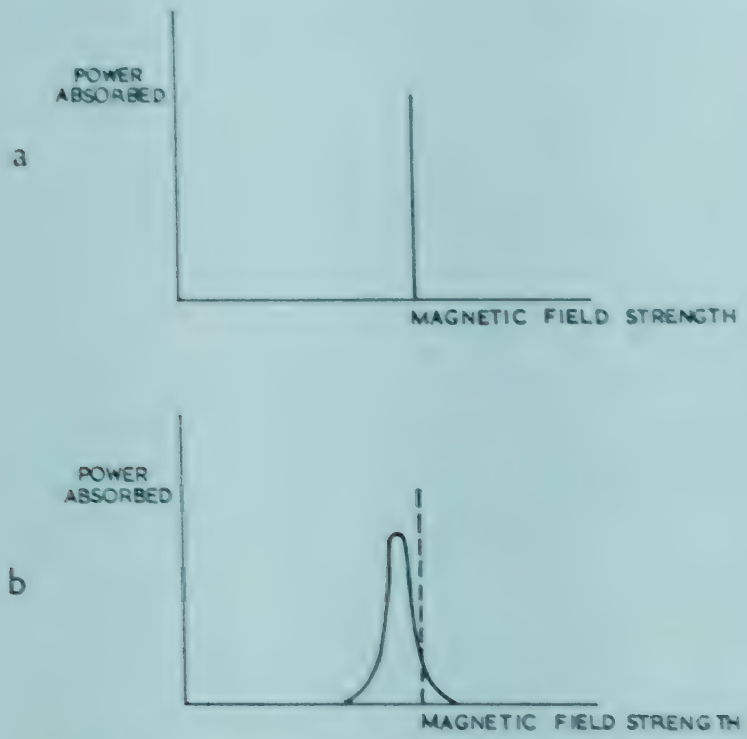


Figure 2. E.S.R. Spectra. 2(*a*) shows the spectrum for a 'free electron'. 2(*b*) shows an actual spectrum, with the free electron spectrum represented by the dotted line.

enough to a magnetic nucleus, the average value of the magnetic field for that electron will be altered and the g-value for resonance will be altered correspondingly. As the hydrogen nucleus (proton) and the nitrogen nucleus are both magnetic, hyperfine splitting is quite common in spectra from biological materials. In some cases the hyperfine splitting is simple and easily understood. Thus the effect of a single proton is to split the line into two equal lines, as shown in Fig. 3. The effect of a nitrogen nucleus, which can take up three different positions with respect

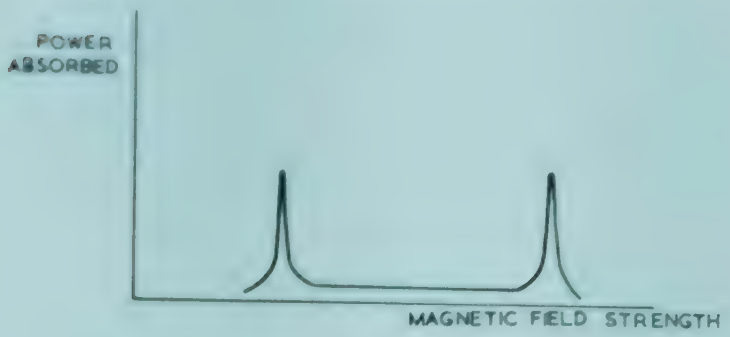


Figure 3. E.S.R. spectrum showing hyperfine splitting by a single proton.

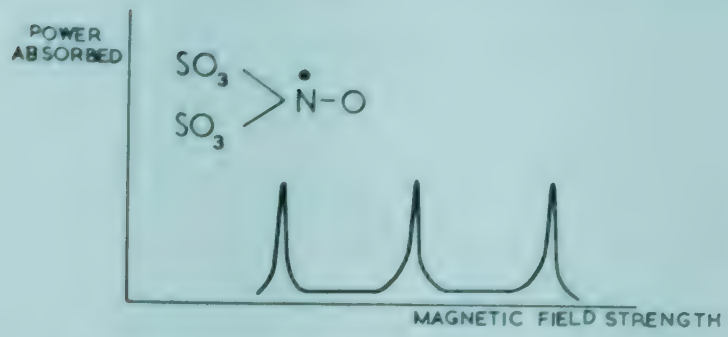


Figure 4. E.S.R. spectrum showing hyperfine splitting by a single nitrogen nucleus. This spectrum is given by the peroxylamine disulphonate ion.

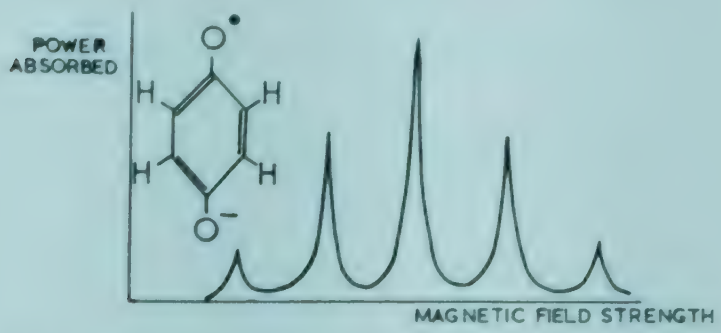


Figure 5. E.S.R. spectrum showing hyperfine splitting by four protons. This spectrum is given by the *p*-benzosemiquinone ion.

to the applied magnetic field, is to split the line into three equal lines, as shown in Fig. 4. If the orbit of the unpaired electron embraces several magnetic nuclei the spectrum can become quite complicated and correspondingly difficult to interpret. A fairly simple case is shown in Fig. 5,

where four protons produce a five-line spectrum with lines of unequal height. The interpretation of a spectrum with hyperfine splitting can often be checked by preparing a specimen in which some nuclei are replaced by isotopes having a different spin. In order to resolve the hyperfine splitting the individual line width must be kept small, and this involves the correct use of the spectroscope as well as the correct preparation of the specimen.

Apparatus

As mentioned earlier, most E.S.R. spectroscopes work at a frequency of about 9,000 megacycles per second. The corresponding magnetic field strength is about 3,000 gauss, which requires a large and heavy magnet. The spectrum can be displayed directly on the screen of an oscilloscope, which is quick and convenient, but greater sensitivity can be obtained by using a pen-recorder. The spectroscope then becomes fairly bulky and

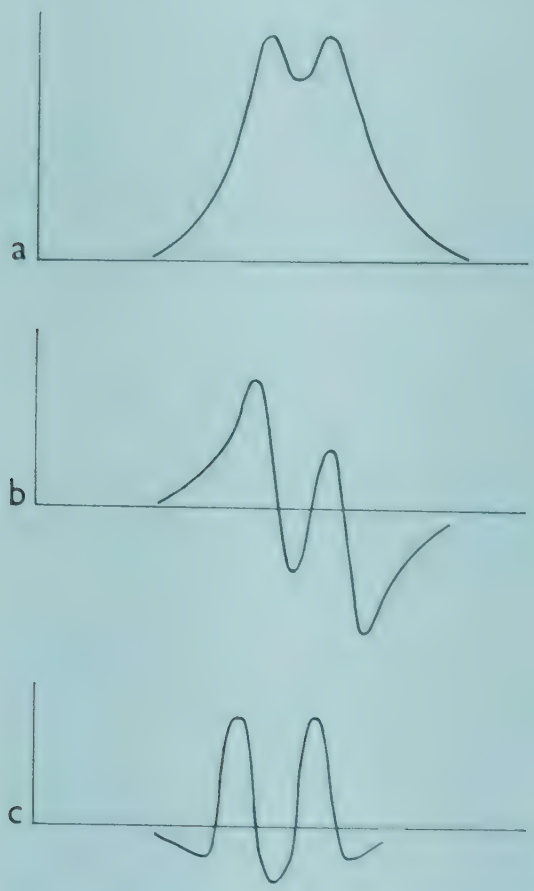


Figure 6. E.S.R. spectra for two closely spaced lines (spacing equal to the line width). (*a*) direct absorption, (*b*) first derivative, (*c*) second derivative.

expensive. Spectroscopes of this type are available commercially. A few spectroscopes have been built to work at lower frequencies. These are cheaper and can be made portable. They are less sensitive, but may not be much less sensitive for biological specimens containing water.

There is considerable flexibility in the choice of the sample. It can be solid, liquid or gaseous, although some liquids, especially water, reduce the sensitivity of the spectroscope. It can also be at any temperature. The size of the sample is usually about 0·5 c.c. in a spectroscope working at 9,000 megacycles per second; it could be considerably larger in a spectroscope working at a lower frequency.

The sensitivity of an E.S.R. spectroscope depends on many factors. As a guide, the most sensitive spectroscopes can detect about 10^{-10} gm. of a solid which has a narrow absorption line. For a liquid sample the concentration has to be about 10^{-8} M, increasing to about 10^{-7} M in the case of water. Most spectroscopes in present-day use are about a hundred times less sensitive than this.

If the spectrum is displayed on an oscilloscope, the direct absorption is usually obtained. The pen-recorder generally gives the first derivative of the absorption spectrum, partly because this comes most easily from the apparatus and partly because hyperfine splitting is shown up better this way. Sometimes the second derivative is recorded, as this gives still better resolution of closely spaced lines. Fig. 6 compares the three types of presentation for the case of two fairly close lines.

Applications to Biological Research

1. STUDY OF X- AND γ-RAY EFFECTS

Although ionizing radiation has been used for more than fifty years to treat a number of diseases, there is still very little knowledge of the way in which the radiation produces the effects it does in living organisms. The aim of a number of workers is to trace, step by step, the pathway leading from the absorption of radiation to the death or modification of the cell. If this could be done it might be possible to make the treatment more successful; at least it would be known whether improvement was possible or not.

The immediate effect of X- or γ-radiation is the production of ions, i.e. molecules with an electric charge due to one or more electrons having been added or removed. (These ions will also be free radicals by our definition.) The details of the initial ionization are reasonably well understood. However the ions are generally very short-lived, and the next step

is the production of free radicals which are somewhat more stable. Although there is agreement on the production of free radicals there is disagreement and even ignorance in many cases about which radicals are produced, and which ones are important for the effect on the cell. This is the information which the E.S.R. spectroscope may be able to provide.

Unfortunately a direct experiment in which the spectrum from living material is obtained during irradiation usually gives no information at all. This is because the lifetime of most free radicals produced by irradiation of wet material is very short, so that the radicals disappear as fast as they are produced and the resultant steady state concentration is too small to be detected. It may be possible, by using the extremely high dose-rates provided by a linear accelerator, to detect free radicals produced during irradiation of wet materials; but the technical problems are very difficult, and no one has reported success so far. Free radicals disappear in liquids because they are free to move about and can combine in pairs to form stable molecules. If the radicals can be prevented from moving, their lifetime will be long, so that a concentration can be built up which can easily be detected in the spectroscope. Stable radicals can be obtained in two ways, either by freezing the material (and it is usually necessary to go down to the temperature of liquid air), or else by using dry material.

Most of the work on the irradiation of biological material has been done by Gordy and his collaborators in the USA. They have irradiated scores of substances in a dry state, and obtained spectra in nearly every case. The substances include amino-acids, peptides, fatty acids, nucleic acids, proteins, enzymes, hormones and vitamins. Other workers have studied carbohydrates in the same way. The situation at present is rather confused, in that a large number of spectra have been obtained, many of which are quite complicated. Some of these cannot be interpreted at all, and none with certainty. Furthermore different workers sometimes give different interpretations of the same spectrum, and occasionally produce different spectra from the same materials. However, a few general conclusions may be drawn. In the case of amino-acids and simple peptides the spectrum usually has a well-resolved hyperfine splitting, characteristic of the particular molecule. Fig. 7 illustrates some characteristic spectra from irradiated amino-acids. It may be possible to identify such radicals from their spectra if they appear in the irradiation of impure biological solids. By contrast with the amino-acids, the proteins give simple spectra without much detail, and the spectra of nearly all irradiated proteins fall into one or two patterns. Fig. 8 illustrates one of these patterns. It is surprising that the more complicated molecule gives such a simple spec-

trum, and not, as one might expect, a combination of spectra for several amino-acids. The protein spectra can be explained by supposing that the unpaired electrons produced by the radiation can move through the protein molecule to a limited extent, until they become trapped at two or three common traps. One such trap is believed to be sulphur.

An example of the use of low temperatures to stabilize free radicals

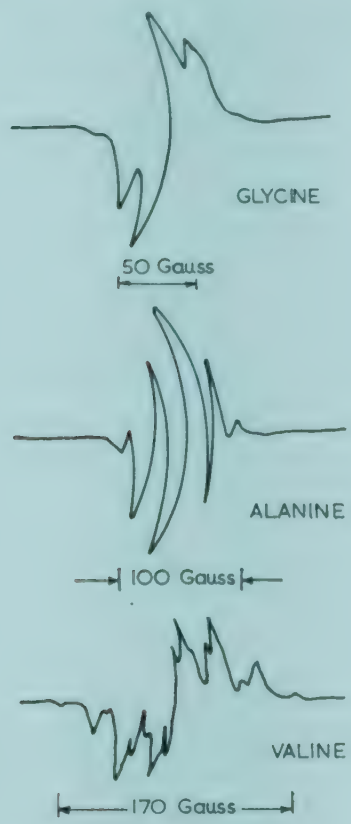


Figure 7. E.S.R. spectra of X-irradiated glycine, alanine and valine. The curves represent first derivatives of the actual absorption lines. After Gordy.

produced by irradiation comes from an experiment by Smaller & Avery, who used E.S.R. spectroscopy to study radiation protection by chemical means. A suspension of yeast in water was mixed with various concentrations of β-Mercaptoethylamine, a substance known to afford some protection against the damage caused by ionizing radiation. Yeast was chosen because it survives freezing to liquid nitrogen temperatures. Samples were quickly frozen, irradiated, and examined in the E.S.R.

spectroscope, still at the temperature of liquid nitrogen. The spectrum was found to have three parts, one due to irradiated water, one to irradiated yeast and one to irradiated β-Mercaptoethylamine. As the concentration of β-Mercaptoethylamine was increased from zero to 5 per cent, the spectrum due to β-Mercaptoethylamine grew in size at the expense of the spectrum due to yeast, which decreased to 20 per cent of its original value. Furthermore, if the temperature of the sample was raised to 125° K., the spectrum due to water disappeared, but the other two spectra were unchanged. These experiments support the view that

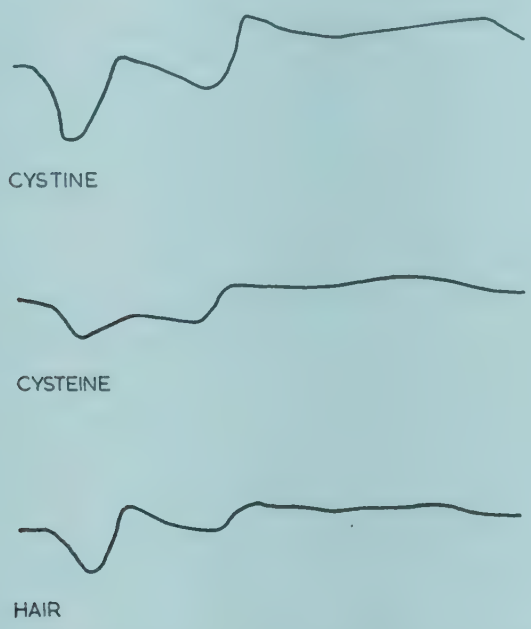


Figure 8. E.S.R. spectra of X-irradiated cystine, cysteine and hair. The curves represent first derivatives of the actual absorption lines. After Gordy.

the protective agent acts directly on constituents of the cell, and not on the radiation products of water.

There are a few examples of living material which are sufficiently dry for free radicals to be stable in them at room temperature. Seeds are one of these, and are particularly useful because the E.S.R. spectrum produced by irradiation can be compared with the effect of the irradiation on the growth of the seeds. Ehrenberg and his collaborators have carried out a series of experiments on seeds, irradiating under different conditions of oxygen tension and water content. The spectrum was always a single broad absorption line, so the only information obtained was the concentration of free radicals and their lifetime. However, a correlation was

found between the concentration of radicals which remained after an initial rapid decay, and the reduction in growth caused by the radiation. The authors are careful to point out that because of the complexity of the living system studied, no definite conclusions can be drawn from this correlation.

II. BIOCHEMICAL REACTIONS

Free radicals have been detected by E.S.R. spectroscopy in a large variety of unirradiated animal and plant materials. The spectra often consist of a single broad line, which does not allow identification of the free radicals. In some cases the spectrum may be an artifact due to the preparation of the sample for the spectroscope, in other cases it may be due to metallic ions or dissolved oxygen. Unless the presence of free radicals fits into a theory which can be checked by other measurements, such E.S.R. measurements are of doubtful value. There are at least two systems in which the E.S.R. measurements fit into a well-established theory. These are photosynthesis in plants, studied by Calvin and Commoner and their co-workers; and oxidation-reduction enzyme reactions. The theory in the latter case is due to Michaelis who proposed that 'all oxidations of organic molecules, although they are bivalent, proceed in two successive univalent steps, the intermediate state being a free radical'.

III. CHEMICAL CARCINOGENS

It has been suggested that chemical carcinogens act through a free radical intermediate. The argument has three parts, firstly that the structure of many carcinogenic molecules is such that they would be expected to form fairly stable free radicals under the right conditions; secondly that free radicals are highly reactive and likely to cause damage and possibly initiate cancer if they come in contact with living cells; and thirdly that ionizing radiation produces both cancer and free radicals. The argument is suggestive rather than convincing. Following up this suggestion it has been shown by E.S.R. spectroscopy that certain aromatic hydrocarbons, including some carcinogens, will produce free radicals if subjected to rather drastic chemical treatment, as attack by concentrated sulphuric acid or by metallic sodium. It would be more convincing if free radicals were produced by a gentler process which might be expected to take place in a living cell. A direct experiment, in which living tissue treated with a carcinogen is placed in the spectroscope, is unlikely to show

anything, because the concentration of free radicals would probably be too small to detect.

An interesting fact, revealed by E.S.R. spectroscopy, is the existence of free radicals in many substances which are known or suspected to cause cancer. They are also present in many substances which are normally regarded as harmless. Fig. 9 illustrates the spectra from tobacco tar, chimney soot, scorched bread and X-irradiated bread. Ordinary bread gave no detectable signal. We have also examined tobacco tar and found a similar E.S.R. spectrum. If the tar was dissolved in benzene and the solution was filtered, a black deposit appeared on the filter paper, which gave the original spectrum; there was no signal from the liquid. We

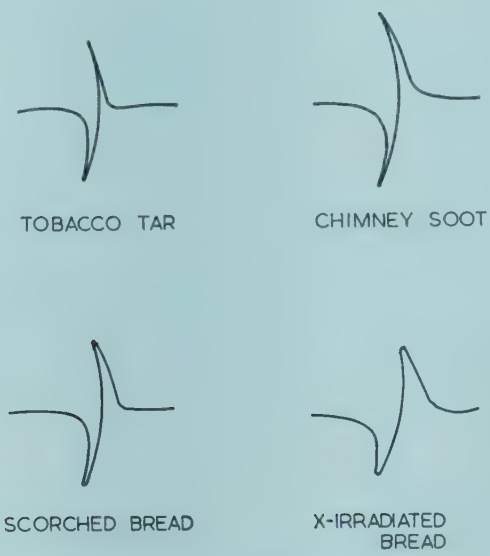


Figure 9. E.S.R. spectra of free radicals in some common substances. After Gordy.

therefore associate the free radical with charred tobacco, so that it will be very similar to the free radicals which can be produced by charring any carbonaceous material, bread for example. This might suggest to some that tobacco smoke was innocuous, to others that burnt toast could cause cancer. Both conclusions might be false.

CONCLUSION

In a relatively short time E.S.R. spectroscopy has revealed the presence of free radicals in a surprising variety of biological material. Some of these revelations may be of doubtful value. The need now seems to be for a more systematic investigation of some of these biological systems,

so that the information obtained from E.S.R. spectroscopy can be fitted in with information from other experimental methods and also supported by theory.

REFERENCES

Calvin, M. (1959) Energy reception and transfer in photosynthesis. *Reviews of Modern Physics*, **31**, 147–156.

Commoner, B., Heise, J. J. & Townsend, J. (1956) Light-induced paramagnetism in chloroplasts. *Proc. National Academy of Sciences*, **42**, 710–718.

Ehrenberg, A. & Ehrenberg, L. (1958) The decay of X-ray induced free radicals in plant seeds and starch. *Arkiv. för Fysik*, **14**, 133–141.

Gordy, W. (1958) Electron spin resonance in the study of radiation damage. *In: Symposium on Information Theory in Biology*. Pergamon Press.

Gordy, W. (1958) Free radicals as a possible cause of mutations and cancer. *In: Symposium on Information Theory in Biology*. Pergamon Press.

Gordy, W. (1955) Microwave spectroscopy of biological substances. I. Paramagnetic resonance in X-irradiated amino acids and proteins. *Proc. National Academy of Sciences*, **41**, 983–996.

Ingram, D. J. E. (1958) *Free Radicals as studied by electron spin resonance*. Butterworths.

Michaelis, L. (1946) Fundamentals of oxidation and reduction. *In: Currents in Biochemical Research*, ed. D. E. Green. Interscience Publishers, New York.

Smaller, B. & Avery, E. C. (1959) Radiation protection and free radicals. *Nature*, **183**, 539–540.

ELECTROENCEPHALOGRAPHY

G. Pampiglione

My task is to present a survey of such study of brain function as can be made with electrical apparatus, and to consider its place in biological research.

Variations in differences of electrical potential occur nearly continuously between regions of the brain and may be recorded through the scalp in man and in other animals with appropriate techniques. The relationship between the waves recorded and the mode of operation of the central nervous system in terms of sensation and its evaluation, or in terms of motor impulses is not fully understood in spite of the large amount of investigation carried out all over the world. Even less is known at both neurological and psychological levels of that supreme function which we call thought and which in our present philosophy is attributed to the brain.

The study of the electrical activity of the brain began a long time ago. Probably the first investigation of this kind on the brain of animals was reported in 1875 by Dr Richard Caton (1875) working in Liverpool and eventually the holder of the first Chair of Physiology in that University. This story can be found in many books and has been put together again a few months ago in a masterly fashion by Lord Cohen of Birkenhead (1959) before the Royal Society of Medicine. Many other physiologists became interested in this new way of approaching the brain, but it was not until thirty years ago that the first investigations on man were published by Hans Berger (1929). He was a wise psychiatrist who had, according to Grey Walter (1953), the reputation of being a crank even amongst his associates. Berger's method was a simple adaptation of the electrocardiographic technique and a potential change of 1/10,000 of a volt could just be detected by his early equipment (a string galvanometer). Later, to his great delight, he acquired a valve amplifier.

In this country the first convincing demonstration of what was then called the Berger's rhythm was given at a meeting of the Physiological Society in Cambridge in 1934 when electrodes attached to Adrian's head

were connected to the Matthews' amplifier and ink-writing oscillograph. In addition to the painstaking studies of Berger and of Adrian & Matthews (1934) some of the early clinical applications of the method were beautifully described by Grey Walter (1936) in his paper, published in the *Lancet*, on the possibility of locating tumours of the brain through the intact scalp. In the United States of America, Gibbs, Davis & Lennox (1935) had reported some peculiar electrical phenomena in epileptics. The war provided great opportunities of studying the effect of brain injury and a paper by Denis Williams (1941) is still relevant. The literature began to grow both in the USA and in Europe and the quality of work in England was then excellent. Speculation about the function of the brain, about the transient electrical storms occurring in epileptics, about the effect of head injuries and about the alteration of the E.E.G. in this or that condition began to multiply. With appropriate evolution of techniques, more reliable and powerful apparatus became commercially available and in the last few years clinical neurophysiology, of which electroencephalography is part, has become an established discipline in many neurological, neurosurgical, paediatric and psychiatric centres in the world.

The growth of these centres has been recently described in a report on electroencephalography from the Royal College of Physicians (1958) based on data collected in 1952–53. In this report it is, however, recognized that in the British Isles, in contrast with some other countries, there has been no commensurate increase in knowledge, and but trivial critical application of the techniques available. In addition to this report the reader might be interested in the recommendations for the organization of the departments of E.E.G. and clinical neurophysiology made by an International Committee on the occasion of the Congress of Neurological Sciences in Brussels (1958).

For those who are not fully conversant with the techniques employed in electroencephalography I shall mention that various regions or points on the scalp have been chosen as more suitable than others to record the activity of the underlying brain. Couples of electrodes on the scalp of the subject are connected to the input of electronic amplifiers. Both spontaneous and evoked variations in potential between any two electrodes are of the order of microvolts. These small signals are amplified a few million times and power is added to activate the recording system. In most commercial sets the 'write out' is obtained by feeding the amplified signals from the brain into the magnetic field in which the recording pen is suspended. The pen moves horizontally on its pivot on a special paper

pulled by a motor at a constant speed. In routine clinical work the speed is usually 1·5–3 cm. per second. The movements of the pen across the moving paper appear therefore as a wavy line. Knowing the speed of the paper it is possible to calculate the duration of the movements of the pen, that is to say the frequency of the recorded waves. Knowing the characteristics of the amplifying circuit, the variations of the differences of potential between any two electrodes can be calculated. The input, the amplifier and the output (recording system) constitute a unit called a channel. For the last ten years or so most E.E.G. apparatus have been

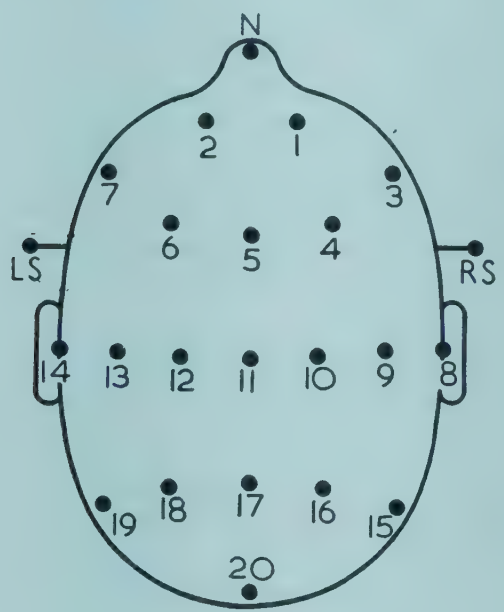


Figure 1. Position of the electrodes on the scalp in routine clinical electroencephalography. (The Hospital for Sick Children, London.)

provided with at least eight channels and often with time markers although equipment with twelve and sixteen channels has been built and occasionally used in England and the USA. In France and other European countries, however, fifteen to twenty or even more channels are very commonly used. Abroad, the introduction of transistors has already greatly altered the size, design and performance of E.E.G. and other polygraphic apparatus.

In routine clinical electroencephalography several electrodes are placed on the scalp in appropriate regions and in our Department we use twenty electrodes stuck to the skin with collodion in the regions indicated in

Fig. 1. When the electrodes are placed directly on the cerebral cortex as during neurosurgical operations the investigation is called electrocorticography. On occasions depth electrodes are used to record the activity of structures beneath the cortical surface. Various types of electrodes may be used to stimulate the brain directly with electrical pulses and to record the effects of such stimulation on the intact and diseased nervous tissue. The electrodes are all connected to the E.E.G. apparatus, but at any one time only some of the possible combinations between them are used. These combinations or 'montages' as they are called are selected by means of switches built in the apparatus. When the combination of the electrodes connected to each pen of the multichannel apparatus is known, the trace of the activity of the various areas explored can be identified. The frequency and amplitude of the waves recorded from a given area of the brain is to a certain extent variable from one normal subject to another. Even in the same person there are considerable differences in relation to the state of alertness, active interest, drowsiness or sleep, as well as differences due to the effect of drugs, age (maturation), and other factors.

It is usually on a somewhat shaky background that any new method of investigation develops and E.E.G. is no exception. It was not long ago that the enthusiasm for this new method of studying the function of the brain suggested to many neurologists and psychiatrists that a grouping of clinical syndromes could be recognized by the shape of the waves recorded and that E.E.G. would make it easy to diagnose and classify any kind of neurological and psychiatric condition. Even the anaesthetists appeared to discover subtle differences in the E.E.G. closely parallel to minor differences in the anaesthetic state. It is now clear that electroencephalography is only one of the many ancillary tools to clinical medicine. Several of the early claims are true only in part; others not at all, but some have stood the test of time and form the basis of our present limited knowledge.

There are various methods of displaying and analysing the electrical activity of the brain. In England, Grey Walter and his school have provided most of the modern technical developments including various types of amplifier, frequency analysers of various kind, topographic (visual and photographic) displays, etc. These tools were created in order to solve particular problems of interpretation and have marked considerable steps forward in the progress of applied neurophysiological work.

I shall now mention some preliminary observations which at the moment are still in the field of research but which in the future might expand the routine applications of electroencephalography. I have been

directly interested in these particular problems and I shall try to follow Rousseau, when he said '*Je sais que la vérité est dans les choses et non dans mon esprit qui les juge: et que, moins je mets du mien dans les jugements que j'en porte, plus je suis sûr d'approcher de la vérité*'.

The first is a rather fundamental point and is related to our method of recording the variations in differences of potential between cerebral regions either directly or through the scalp. The range of new observations that one can make on everyday clinical material is limited by the performance of the commercially available apparatus. Very little attention has been paid until recently to the slowly changing potential differences within the brain. The response of an RC coupled amplifier to slow frequencies is determined by its so-called time constant. This in most

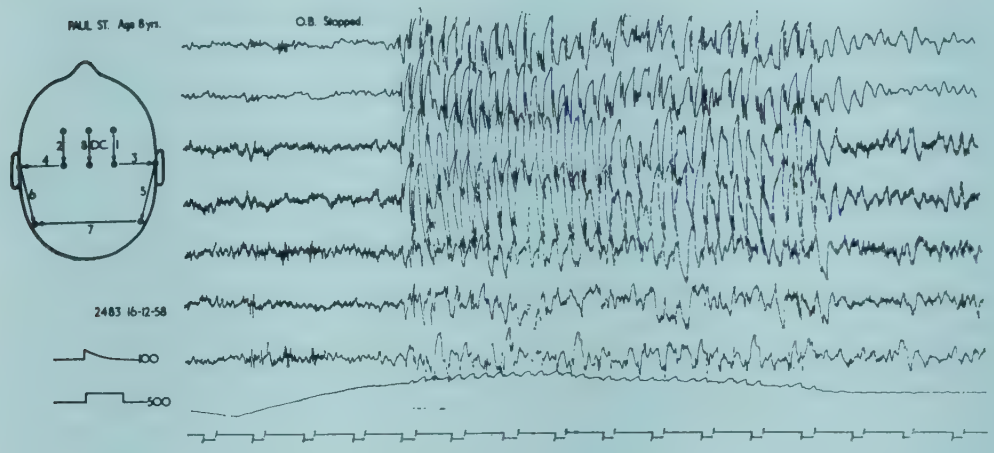


Figure 2. Conventional E.E.G. with 0·35 seconds Time-constant (Channels 1–7) and record obtained through a d.c. amplifier (Channel 8) during a short-lived seizure. Notice the slow potential change preceding the attack.

commercial apparatus is one second at the most and no data could be collected by most workers on slower phenomena. Neurophysiologists however have become increasingly interested both in much slower potential changes and in steady potential differences, for example those detectable between various cortical regions as well as parts of the brain-stem and thalamus. With suitable types of amplifier we may be able to measure slowly changing potential differences in the brain of man. An indication that the 'setting' of one or more regions of the brain is slowly changing with respect to others may become of considerable theoretical and practical importance, as these phenomena possibly underlie processes of excitation and inhibition. During some seizures in man, slow changes

of potential occur in various regions of the brain often preceding by an appreciable interval the beginning and end of the clinical phenomenon. For the last two years I have been collecting data on children during some 200 seizures of various kinds (Pampiglione, 1959). The record, obtained with a directly coupled amplifier, appears somewhat different from the ordinary type of E.E.G. recorded with a relatively short time constant. Fig. 2 illustrates what may happen about the time of a short-lived seizure.

Let us now divert our attention from the record of what happens 'spontaneously' in the brain to the features obtained by stimulating the subject. A large amount of literature has accumulated on this matter but there is room for further observations. When we are called by name, a number of phenomena follow, some of which might be detected in the electroencephalogram. Our reaction to being called, however, depends on several factors (Pampiglione & Ackner, 1958). Amongst the first aspects to consider is our ability to hear the call and also to recognize its meaning. In a particular set of circumstances one modality of stimulation might be much more effective than another.

In a study of the normal cerebral maturation in the dog that I am carrying out at the Human Nutrition Unit of the M.R.C. at Mill Hill with Prof. B. S. Platt, it appears (Pampiglione, Platt & Stewart, 1959) that E.E.G. changes in a pup in response to loud noises are minimal in the first two or three weeks of life both during the waking state and sleep. The E.E.G. changes in response to an olfactory stimulus, on the other hand, are very well developed even in the first week of life. Later, however, the response to a whistle or to a hand clap becomes very prompt, while the response to the smell of food is slower and does not persist, particularly if the dog has been recently fed. In young babies, although the motor responses to a loud noise might be considerable, the E.E.G. changes are small. The E.E.G. responses to an auditory stimulus become more obvious after the age of six to nine months. In the study of deafness in children it is often difficult to assess audiometrically the type of either partial or total hearing loss. In general, very little, if any, use has been made of electroencephalography as a means of differentiating various types of deafness. The information obtained by this method is not easy to interpret and requires a considerable amount of experience to avoid common pitfalls, but is potentially very useful. The E.E.G. change that may occur after an auditory stimulus is not a specific response of the auditory cortex, but is a less specific change of a kind that might follow other modalities of stimulation. It only means 'some signal has arrived', not that the stimulus has been 'recognized'. In a group of children considered deaf on

audiometric tests there have been a number who showed well-recognizable E.E.G. responses following a noise both at low and high frequencies. My interpretation is that these children have a fairly well-preserved middle and inner ear, and a nerve that conducts to the brain stem. From there the irradiated E.E.G. response must take place. Such deafness could be regarded as of central type; that is to say the child is unable to integrate and to interpret the signals that reach his brain stem and is therefore unable to understand their meaning, appearing 'deaf' on audiometric tests.

We shall now consider a third application of electroencephalography of some interest to surgeons. For nearly three years at The Hospital for Sick Children we have been taking E.E.G's in the operating theatre during various cardiothoracic interventions. I am indebted to Mr David Waterston the cardiothoracic surgeon who not only tolerated but appeared even to welcome within the operating theatre the somewhat bulky tools of our trade.

The arrangements in the operating theatre have become standardized. The electrodes are stuck to the scalp with collodion in predetermined positions to allow a comparison between the records taken during the operation and the pre- and post-operative E.E.G's. The electrodes recording the cardiogram were placed over the right and over the left shoulder. The E.E.G's were taken continuously for several hours throughout the operation from the early phases of induction of anaesthesia to the closing of the skin. This somewhat laborious routine was rewarding in that a number of unsuspected events could be documented.

During various operations on the heart and main vessels the blood supply to the brain as well as its venous return may have to be disturbed and often the cerebral circulation has to be occluded completely for short periods. The interference with the cerebral circulation may last several minutes under moderate hypothermia. In addition, during the intervention, and sometimes even before this, the heart action may become insufficient, or stop altogether, independently, as it seems, of the wish of either the operator or the anaesthetist (Fig. 3). Cardiac massage might have to be carried out sometimes for a quarter of an hour or longer before the heart beat is re-established. Nearly two years ago, with Waterston (Pampiglione & Waterston, 1958), we described some of our preliminary findings before the E.E.G. Society, but I shall mention now some of the unpublished features mixing freely both clinical and E.E.G. observations.

When the flow in the superior vena cava was obstructed a congestion

of the face was often noticed accompanied by an increase in the size of the pupils. There was, however, no appearance of muscle activity either clinically or electrographically for occlusions lasting up to 40 seconds. In the E.E.G. a diminution in fast activity was often noticed in about 20 seconds, together with the appearance of slow waves in variable amounts. This change was rapidly reversible after re-establishing the circulation. When the pressure in the superior vena cava was measured

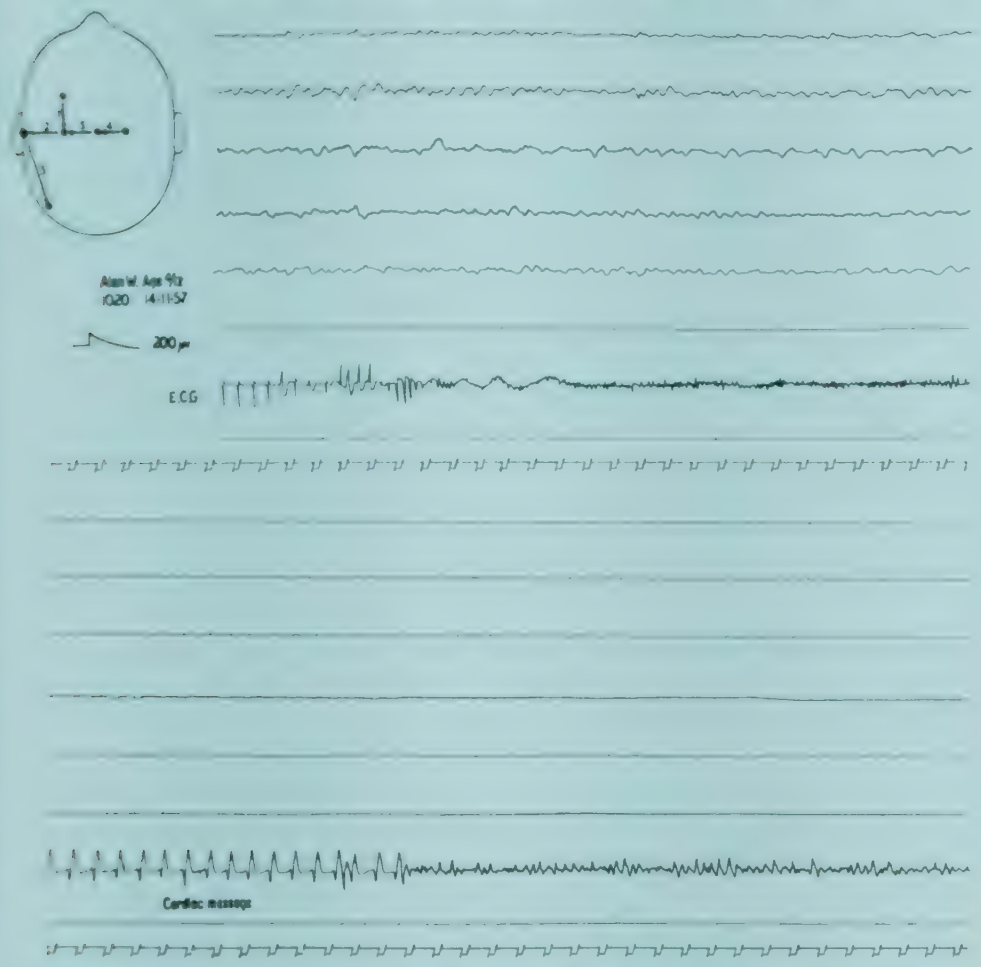


Figure 3. During a cardiothoracic exploration the heart action suddenly became extremely poor. In addition to the electrocardiographic changes, gradually the brain waves became slower and smaller. Soon the traces from the brain became flat in spite of the cardiac massage.

it often showed a rapid increase reaching a plateau in about 15 seconds probably related to the arterial pressure. At this level, definite oscillations in the raised venous pressure appeared synchronous with the pulse. At

about the same time in the E.E.G. the slow wave activity appeared both at normal temperatures and during moderate hypothermia. When there was only a small rise in venous pressure, the cyanosis was not severe and no change occurred in the E.E.G. In this case it was often found that an

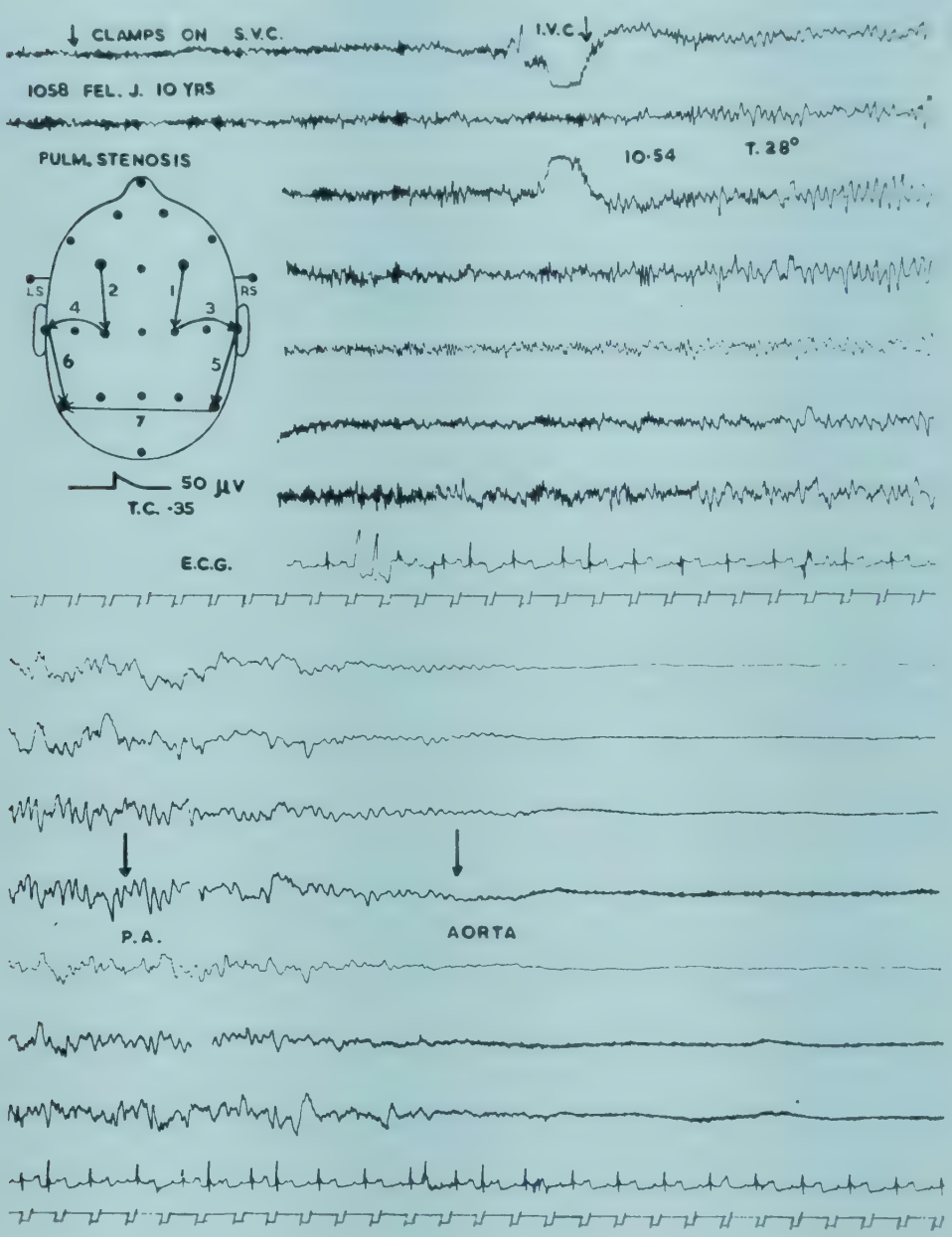


Figure 4. The usual E.E.G. changes during occlusion of venous flow and subsequent absence of cardiac output. Moderate hypothermia (28° C.). After the appearance of slow activity the traces became flat and some muscle action potentials appeared from the temporal and temporo-occipital regions.

abnormal venous channel or a large azygos vein was present above the clamp, thus preventing a full obstruction of the venous drainage from the brain.

When, in addition to the superior vena cava, the inferior vena cava was also occluded, as well as the venous congestion there was also a lack

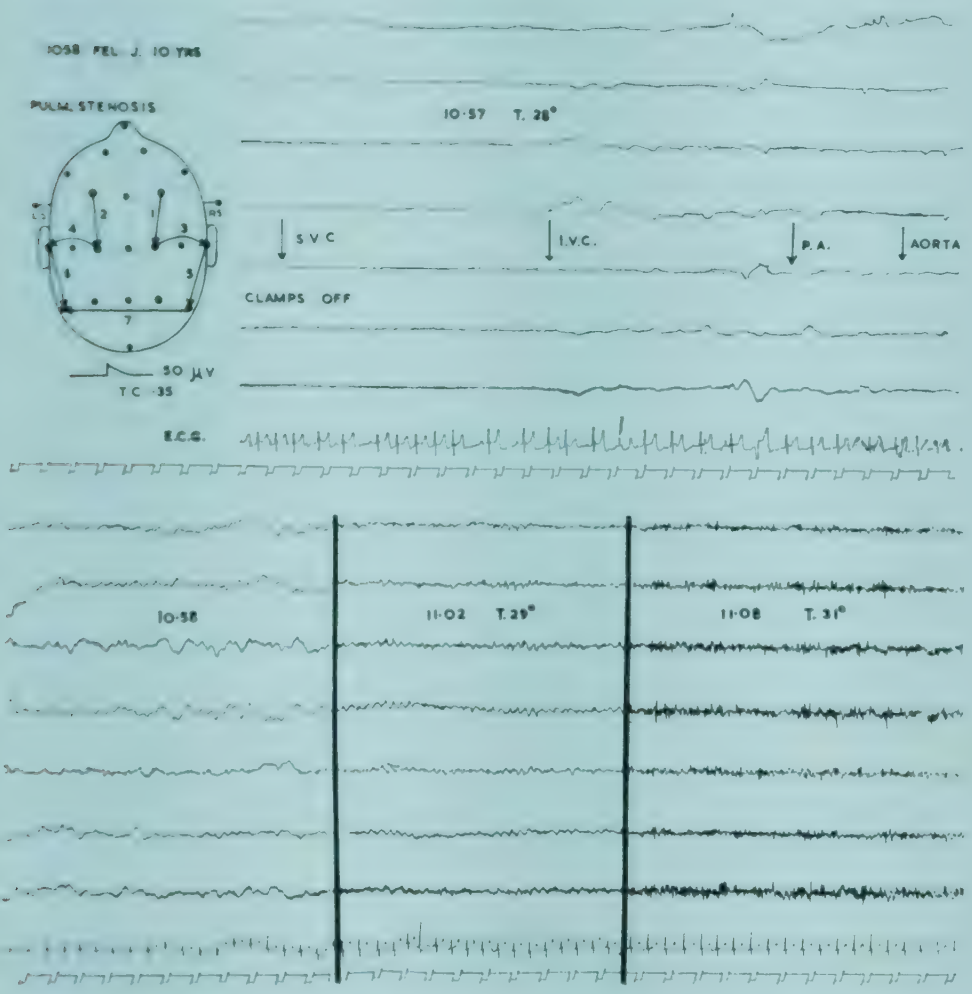


Figure 5. The same patient during recovery following the re-establishment of circulation. The time intervals are marked in each section. Both clinical and E.E.G. recovery are rapid for short periods of obstruction to the cerebral circulation.

of supply of fresh blood to the brain, after the exhaustion of the pulmonary circulation. In these cases, as well as after the clamping of the aorta, there was at first an increase in slow waves in the E.E.G. while the patient became congested and cyanotic. The E.E.G. after the increase in slow waves became flat, but soon muscle action potentials appeared both from

the scalp and from the shoulders. Clinically there was occasionally, in about half a minute, a tonic spasm of the muscles of the head and neck and often of the arms in spite of curarization. The pupils became extremely widely dilated and then the patient might slowly relax. The E.E.G. remained completely flat during the cerebral ischaemia (Figs. 4 and 5).

After re-establishment of the circulation the patients remained fully relaxed and, in a few minutes, the dilated pupils became smaller. In the E.E.G. large runs of very slow waves separated by intervals of relative equipotentiality were seen similar to those often recorded after prolonged spontaneous or electrically-induced seizures. The flattening of the E.E.G. soon after exclusion of cerebral blood flow always outlasted the time of occlusion of the vessels. On occasions the record remained completely flat for over half an hour before the recovery. Sometimes in the absence of any heart beat and of any respiratory effort with an apparently absent electrical activity from both brain and heart it was difficult to decide whether the patient should be considered still alive, and as yet I have found no reliable criteria, in the first hour, for encouraging the surgeons or the anaesthetists to continue in their resuscitation efforts or to abandon them. In three of the patients who had been apparently dead for over half an hour, there was a complete recovery with full intellectual, motor and sensory function by the next day, as well as an E.E.G. pattern of nearly full recovery.

Such problems present only a glimpse of the formidable fortress that we are trying to penetrate. This fortress with an incredible number of secret passages, booby traps and labyrinths is what is often called cerebral function. It appears surrounded by an enormous, poorly illuminated borderland where many research workers get lost, ingenious techniques become side-tracked and strategic plans are often defeated, but some get through.

REFERENCES

Adrian, E. D. & Matthews, B. H. C. (1934) Berger's rhythm. Potential changes from the occipital lobes in man. *Brain*, **57**, 355.

Berger, H. (1929) Uber Das Elektrenkephalogram Des Menschen. *Arch. Psychiat. Nervenkr.*, **87**, 527.

Caton, R. (1875) The electric currents of the brain. *Brit. Med. J.*, **2**, 278.

Cohen of Birkenhead (1959) Richard Caton (1842–1926): Pioneer electrophysiologist. *Proc. R.S.M.*, **52**, 645.

Gibbs, F. A., Davis, H. & Lennox, W. G. (1935) The E.E.G. in epilepsy and in conditions of impaired consciousness. *Arch. Neurol. Psychiat., Chicago*, **34**, 1133.

International Federation of Societies for E.E.G. and Clin. Neurophysiol (1958) Proceedings of the General Assembly, IVth Internat. E.E.G. Congress. Bruxelles 1957. *E.E.G. Clin. Neurophysiol.*, **10**, 367.

Pampiglione, G. (1959) Slow potential changes during seizures. *Proc. Internat. Conf. Med. Electronics.* Unesco, Paris (in Press).

PAMPIGLIONE, G. & ACKNER, B. (1958) The effect of repeated stimuli upon E.E.G. and vasomotor activity during sleep in man. *Brain*, **81**, 64.

PAMPIGLIONE, G., PLATT, B. S. & STEWART, R. J. C. (1959) Long-term E.E.G. and E.C.G. studies in young rats, pigs, and dogs. *E.E.G. and Clin. Neurophysiol.*, **11**, 603.

PAMPIGLIONE, G. & WATERSTON, D. (1958) Preliminary E.E.G. observations during partial and complete occlusion of cerebral blood flow. *E.E.G. Clin. Neurophysiol.*, **10**, 354.

ROYAL COLLEGE OF PHYSICIANS LONDON (1958) *Report of Committee on Electroencephalography*

WALTER, G. W. (1936) The location of cerebral tumours by electroencephalography. *Lancet*, **II**, 305.

WALTER, G. W. (1953) *The Living Brain*. Duckworth, London.

WILLIAMS, D. (1941) The electroencephalogram in acute head injuries. *J. Neurol. Psychiat.*, **4** (N.S.), 107.

PAPER CHROMATOGRAPHY

IVOR SMITH

The technique to be described is a comparatively new one, the first paper on this subject being published in 1944 and the first in the biochemical-clinical field in 1947. Since that time the number of papers mentioning this technique has gone up in leaps and bounds and there is now almost no branch of the chemical or biological sciences where paper chromatography has not contributed to the increase in knowledge and understanding of the subject. The reasons for this are quite plain. Paper chromatography is simple and quick, it requires only minute amounts of substances and, perhaps most important, it requires no expensive or delicate apparatus.

Paper chromatography is a technique for separating microgram amounts of closely related substances and a paper chromatogram is prepared in the following way. A drop of the solution containing the compounds to be separated is placed some 3 or 4 cm. from the edge of a strip of paper and allowed to dry. The strip is then placed so that some 2–3 mm. dips into a solvent which commences to flow along the strip, over the dry spot and towards the far end of the paper. When the solvent has moved some satisfactory distance, the strip is removed from the apparatus and dried. In this way a one-way chromatogram is prepared and the compounds present in the original mixture are found to be completely or partially separated along the length of the strip. However when a mixture contains many substances of interest the partial separation so achieved is insufficient to allow of the identification of the components and so recourse is had to two-way chromatography.

A two-dimensional chromatogram is prepared in a similar manner. A drop of the solution is placed near the corner of a square or rectangular sheet of paper and solvent is allowed to flow along the paper as for a one-way chromatogram, thus separating the compounds along a line near, and parallel to, one edge of the paper. The sheet is then dried and turned at right angles and a second solvent is allowed to run across the paper; thus the partially separated compounds serve as a number of much

simpler mixtures for separation in the second direction. The second solvent is chosen such that its properties are quite different from those of the first solvent so that substances which run together in the first solvent will separate in the second.

Having separated the substances of interest the problem of locating them still remains. The most widely used method is to dip the dry chromatogram through a solution of a chemical reagent which reacts with these substances with the production of a colour. The location reagent may react with only one substance or with many substances on the chromatogram but, by a judicious choice of reagents, it is possible to apply a number of these in sequence thereby obtaining a great deal more information than could be obtained with any one reagent alone.

In order to define the position to which a substance moves on the chromatogram the term R_f was introduced, thus:

$$R_f = \frac{\text{the distance the substance moves from the origin}}{\text{the distance the solvent front moves from the origin}}$$

The R_f of a pure compound, therefore, is a physical constant just as much as its melting point is. Nevertheless, just as the melting point of a pure substance is altered in the presence of other compounds so the R_f of a pure substance can be affected by the presence of other substances in a mixture. Because of this, and particularly where one is seeking the presence of unusual spots on the chromatogram, it is usually much more satisfactory to familiarize oneself with the normal 'map' or pattern and any variation from this is then immediately obvious.

It cannot be too greatly stressed that it is often essential to run two-way chromatograms, even with different pairs of solvents, and to make the most use of the location reagents available. All of the above points are illustrated in Fig. 1.

In practice, chromatography is usually carried out with ascending or descending solvent flow in an air-tight glass tank. In the former method the mixture is placed some 3 cm. above the bottom of a sheet of paper which is held on a metal or polythene frame (see Fig. 2). The frame is then placed in a tray containing solvent which immediately begins to rise up the paper by capillary flow. In this particular apparatus the sheets are 10 in. square and each can be used to run ten one-way chromatograms or one two-way chromatogram; thus fifty one-way or five two-way chromatograms can be prepared at one time. Although an enormous amount of work of clinical and biochemical interest has been done in this apparatus, papers and tanks of all sizes have been used. Indeed, all

that is required to run a satisfactory chromatogram is an air-tight tank. The advantages of this frame technique in the apparatus shown are: chromatograms prepared on different occasions can be directly compared

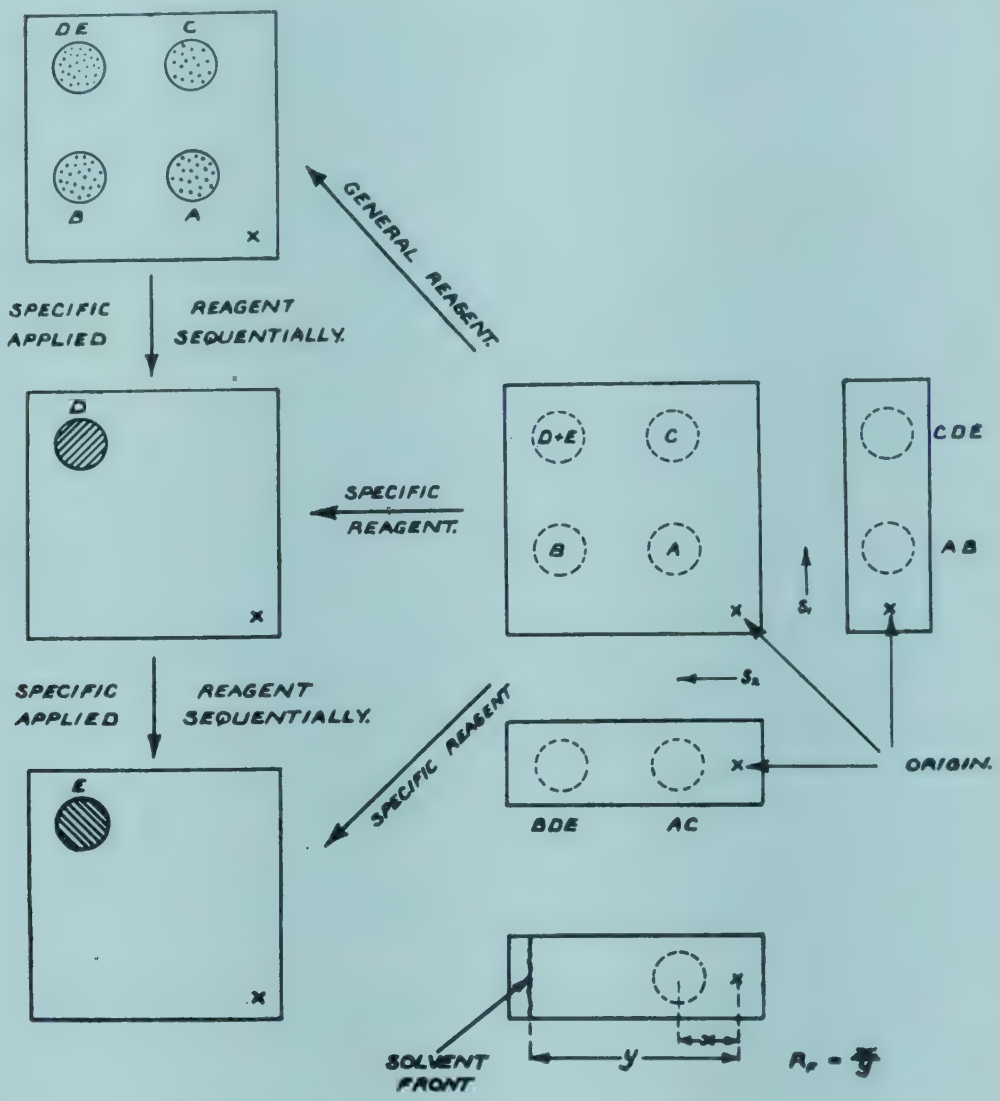
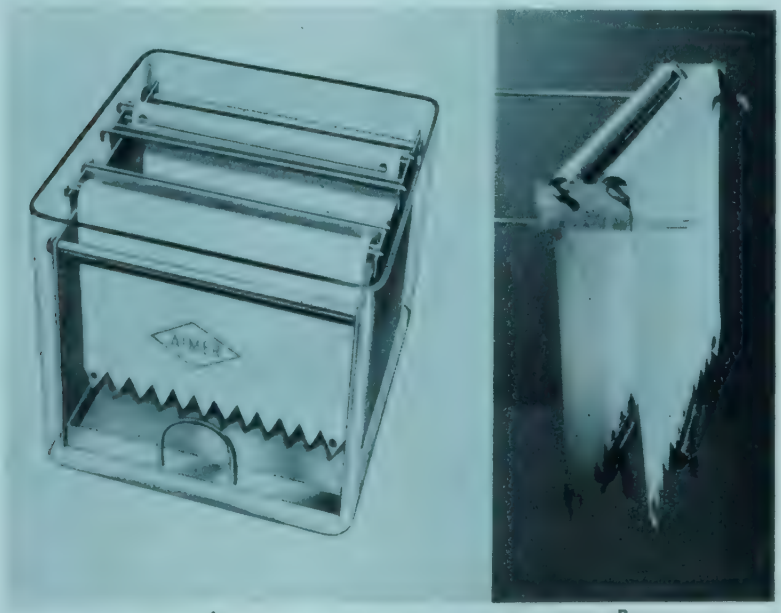


Figure 1. Diagram illustrating one- and two-dimensional paper chromatography, the use of general and specific reagents and sequential application of reagents. ABCDE are five components present in a mixture. In solvent S_1 the R_f values of A and B are equal; those of C, D and E are also equal. In solvent S_2 the R_f values of A and C are equal; those of B, D and E are also equal. A one-way chromatogram using either solvent therefore shows only two spots on location with a general reagent. A two-way chromatogram shows four spots on location with a general reagent, but by the application of specific reagents to duplicate chromatograms one of the spots is shown to be a mixture of two components D and E having the same R_f in both solvents. When the reagents can be applied sequentially only one chromatogram need be prepared, otherwise three separate chromatograms are necessary.

A B

Figure 2. *a.* The 'Universal Apparatus' is assembled for ascending chromatography. The frame, holding from one to five papers, sits in the solvent in the bottom of the tray. *b.* The frame is being assembled. The papers are placed horizontally on the frame and are held apart by collars; finally the end plate is placed in position and held by the nuts. The solutions to be chromatographed are applied to the origins with the frame still in the horizontal position shown.



A B

Figure 3. *a.* The apparatus is assembled for descending chromatography. The rear trough is tilted forward to illustrate how it is held firmly on the frame end plates and also the method of suspending the papers in the trough. *b.* Drying the papers. The trough is removed from the tank and suspended over parallel bars, any residual solvent being pipetted out to prevent further solvent flow. Clips are attached to the lower ends of the papers to prevent them flapping together.

because of the very high order of reproducibility obtainable; the wet papers, which have low mechanical strength, need never be handled as the frame can be picked up directly; the whole can be easily dried by placing in front of an air-blower or fan.

Paper chromatograms can also be prepared by descending solvent flow. In this procedure the paper is hung from a trough containing solvent which then flows down the freely hanging paper, over the spot and on towards the bottom of the sheet. Both one- and two-way chromatograms can be prepared and the separations obtained are similar to those found with the ascending technique. In the author's opinion the ascending technique is far preferable to the descending method as it is much simpler to practise and it results in much more reproducible results.

The choice of a solvent for chromatographic separation is still a matter of intelligent guesswork or trial and error. However some simple rules are available. If a substance is very soluble in a solvent then it will run with the solvent front on a chromatogram; conversely, if it is insoluble it will remain at the origin. Thus by mixing the two solvents in different ratios the substance to be chromatographed can be made to travel to any position on the paper. When we consider the separation of any two substances, it is obvious that the more close they are in chemical constitution then the closer will be their R_f values and the more difficult it will be to separate them. A similar argument applies when the solvent is composed of three or four components. It follows, therefore, that the more complex the original mixture, the more difficult will it be to find solvents which will adequately separate the substances present. Amongst the other factors which affect the separation of the components of a mixture are the temperature, grade of paper, ionization of the substance in the solvent, other components with similar R_f values, etc.

TABLE I. *The variation of R_f in binary solvent mixtures. The R_f values of the phenylthiohydantoins of three amino-acids are shown. All are insoluble in heptane (H) and therefore remain at the origin; all travel with the front in pyridine (P). By varying the composition of the mixture, each derivative can be made to run to any part of the paper desired*

	Solvent mixtures					
H	100	80	70	60	50	0
P	0	20	30	40	50	100
Lys. PTH	0	10	22	60	100	100
Ala. PTH	0	38	44	82	100	100
Leu. PTH	0	77	82	94	100	100

TABLE II. *The composition and properties of some solvents useful for the separation of various families of compounds*

Solvent for	Composition by volume		Preparation	Length of run	Use
Amino-acids	Butanol	60	As required	Overnight	First solvent.
	Acetic acid	15			
	Water	25			
	Phenol	80	In bulk	,,	Second solvent.
	Water	20			
	NH_3.88	0·5			
Indoles	*iso*-propanol	200	As required	,,	First solvent.
	Water	20			
	Ammonia.88	10			
	Butanol	60	,,	6 hrs.	Second solvent.
	Acetic acid	15			
	Water	25			
Sugars	Ethyl acetate	60	,,	Overnight	First—one-way only required.
	Pyridine	25			
	Water	20			
Steroids, C17, O_2	Water	10	As required or in bulk	3 hr. run after equilibration at 37° C.	Use upper phase of two-phase mixture. One-way only required.
	Methanol	40			
	Petrol	50			
Steroids, C21, O_5	Water	250	,,	,,	,,
	Methanol	250			
	Benzene	500			

Location reagents are prepared in solvents in which the substances of interest are not soluble. The reagent is poured into a tray of suitable dimensions and the dry chromatogram is dipped rapidly and evenly through it with the result that a very even distribution of reagent is obtained over the whole paper and that the colours produced by different substances in different areas of the paper will be quantitatively comparable. The colours may appear at once or after a few hours on standing at room temperature, or the paper may require heating at up to 110° C. for some minutes before the colour reaction occurs. The information obtained is then noted and a tracing made or a picture taken, or the spots may be delineated with pencil on the actual chromatogram; this delineation is essential if further reagents are to be applied. After each successive reagent is applied the further information obtained is also recorded; in some cases up to four reagents can be applied in sequence. Apart from the obvious advantage of obtaining extra information, this multiple dipping technique possesses the following other advantages. First, a large number of specimens can be rapidly screened for a wide

range of chemical families, second, chemical and metabolic changes involving the loss or alteration of a functional group from a molecule can be followed; third, two overlapping compounds with different functional groups can be differentiated from one compound with dual function as, in the former case, the colours will only overlap but in the latter case, they will be completely coincident.

The methods described lend themselves equally to qualitative or quantitative analysis. In practice, however, it is much simpler to use one-dimensional procedures for routine quantitative analysis if this is possible. The substance to be determined can be measured either directly from the colour intensity on the paper, or by eluting the colour into some solution, or by eluting the unreacted substance into solution and then carrying out the colour reaction. In the first case, the mixture is run in parallel with known amounts of the pure substance and from the colour intensities of the standards the quantity present in the mixture is gauged either visually or by means of a densitometric scanning. In the second case the colours produced by the mixture and the standards are separately eluted into a solvent which is then made up to a standard volume and the colour intensity is determined in a colorimeter. In the third case, the mixture is applied to two origins on the same paper so that they run in parallel. One of these runs is treated with reagent and then realigned with the untreated strip when the position of a particular compound can be ascertained. The section of paper containing the substance is then cut from the remainder of the strip, dissolved off into a suitable solvent and the colour reaction is carried out in this solution, the colour produced again being measured in a colorimeter.

The chromatographic separation procedure described is a simple analytical one, that is, it reveals the position of all substances which react with a given reagent. However, during the course of an investigation it is often found that a given spot cannot be identified as a known compound and it then becomes necessary to recover larger amounts of the substance in order to determine its chemical nature. Paper chromatography is equally well suited to this demand as a sheet 10 in. wide will easily carry 50 mg. of mixture; this fact is not at all well appreciated by many workers in this field.

A streak chromatogram (Fig. 4) is prepared by streaking the solution from a pipette along the length of the origin, this being equivalent to many separate drops applied simultaneously. The streak is allowed to dry and the process repeated until a sufficient amount has been applied. The chromatogram is then run in the normal way and dried again. Guide

strips are cut from each side of the paper and treated with reagent to locate the substances concerned and to show the position of the new substance. The strips are realigned with the untreated paper, the position of the new substance is marked, and the section of the paper containing it is cut out. The strip so obtained is joined to a clean piece of paper and set up as a descending chromatogram. The solvent chosen is one in which the substance is completely soluble and so it travels with the front and drips off into a small test-tube held just below the lower end of the strip. This washing-off process, known as elution, is usually quantitative. If needs be, the whole process can be repeated with a different solvent until a single pure substance is obtained in the eluate. The pure material can then be investigated by standard chemical procedures.

Before embarking on paper chromatography it is nearly always

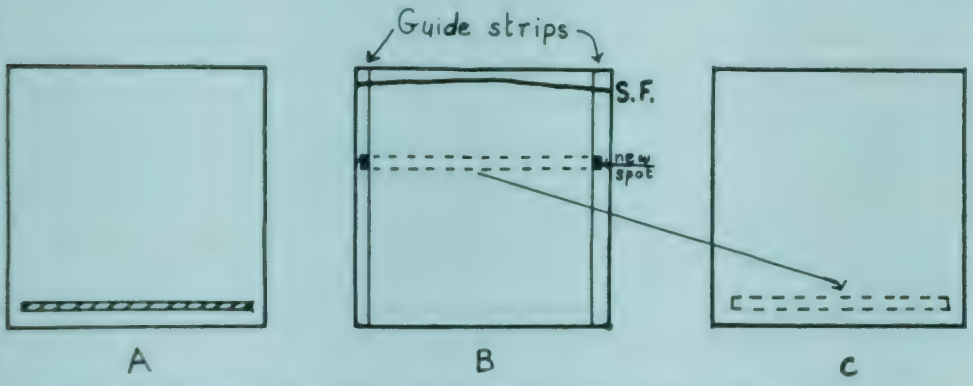


Figure 4. A. A solution is 'streaked' on to the chromatogram which is then run in the normal way in solvent S_1. B. The dried chromatogram is taken, guide strips are cut from it and located with a suitable reagent. The strips are re-aligned with the untreated portion of the chromatogram, the position of the new spot is marked with a pencil and the strip is cut out. C. The substance is eluted from the untreated strip and applied to a second sheet for chromatography in solvent S_2. The process is repeated until a pure substance is separated. (S.F. = solvent front.)

essential to carry out some preliminary purification or concentration of the substance or substances of interest. Inorganic salts almost invariably interfere with solvents most useful for the separation of organic compounds and must therefore be eliminated. In the case of those substances present in greatest amounts in normal and abnormal urines (amino-acids, sugars) this may be the only treatment necessary; de-salting is brought about either electrolytically or by ion exchange procedures. However, when one has to deal with compounds present in micro amounts (steroids, barbiturates) then it becomes necessary to remove not only the salts but also most of the organic compounds present in order to avoid over-

loading the paper. Fortunately many of these latter compounds are relatively non-polar and can be extracted into organic solvents which eliminate both the salts and the polar organic substances simultaneously. At this stage the sample is usually sufficiently pure to apply to the chromatogram. The quantity to be applied will depend on the individual compounds sought and the sensitivity of the location reagent. The reagent most generally used for amino-acids will detect 2–3 μg. of any amino-acid and therefore 20 μl. of de-salted urine will be sufficient for a one-way chromatogram when using a 10-in. square paper; roughly three times the quantity is used for a two-way run. For sugars it is necessary to apply some 15 μg. of each sugar which may be present in up to 50 μl. of de-salted urine. For cholesterol some 250 μg. must be applied and this is usually obtained by ether-ethanol extraction of 0·2 ml. serum. For the steroids some 10–20 μg. of each must be applied, but this amount is found in about 150 ml. urine (one-tenth day's urine). This quantity of urine must therefore be extracted and concentrated so that the whole can be applied to the origin. In fact the extract obtained is too impure to yield satisfactory chromatograms and a further column separation is often necessary before satisfactory paper chromatograms can be obtained.

The applications of paper chromatography in biochemistry can be considered under the following headings:

1. The diagnosis of disease including the differentiation of diseases presenting with very similar clinical symptoms.

2. Determination of the nature of ingested or injected foreign bodies and their metabolites, e.g. drugs, dope, etc.

3. The evaluation and control of therapy.

4. Determination of normal and abnormal metabolic pathways, often with the aid of radio-isotopes.

5. Determination of the chemical structure of biologically active molecules.

The diagnosis or confirmation of many diseases has, with the advent of paper chromatography, become more a question of a quick routine chromatographic analysis than of what was often a lengthy clinical investigation; this is particularly true of many conditions which are genetically controlled. In discussing this it will be most convenient to consider first those families which require one-dimensional runs only and then to consider those which require two-way runs.

Normal urine contains no sugar or, to put it more carefully, when 50 μl. of de-salted urine is run on a one-way chromatogram no sugar can

be detected with the usual reagents. Any urine which reacts positively must, therefore, be abnormal; pregnancy urine often contains lactose and glucose but it is debatable whether this is abnormal. Apart from this some seven conditions are known where sugar is found in the urine. It is rare to find more than two sugars associated with a particular condition and so the determination of the nature of the sugar is very simple; the clinical

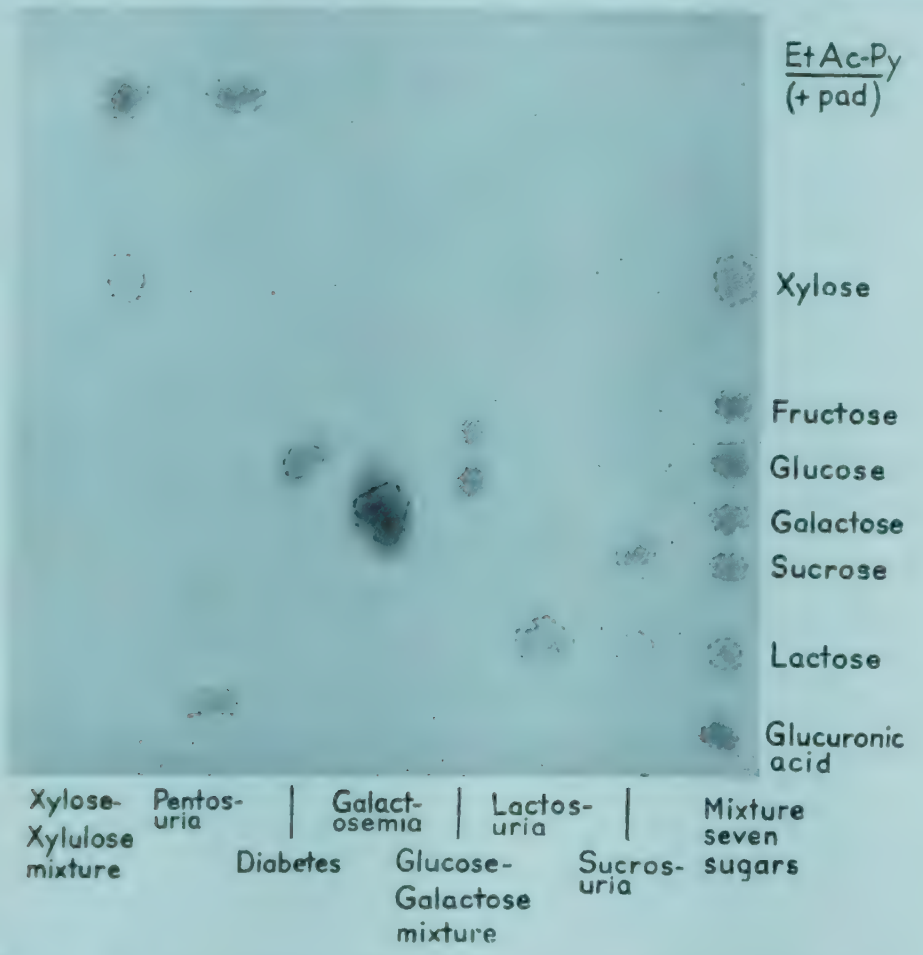


Figure 5. The picture illustrates the separation of various marker mixtures run in parallel with normal and abnormal glycosurias. Sugars as similar as glucose and galactose are easily separable.

symptoms will either suggest or confirm the identity of the sugar.

In this 'anxiety' age, the number of persons having recourse to barbiturate sedation runs into many thousands. Indeed, it has been estimated that about 5,000 cases of barbiturate poisoning are investigated each year. This is not surprising as many of those being tranquillized are precisely the ones with suicidal tendencies. Furthermore there are the 'knock-out

drop' offences against humans and racing animals as well as murders. Thus the forensic chemist may be required to examine tablets, stomach contents, the dregs from a cup or glass, the stains on a sheet and so on for the presence of barbiturates or other drugs. This problem may be particularly difficult when the ingested drug is rapidly metabolized as different quantitative patterns may be found depending on how long the drug has been in the body. The problem is rendered even more difficult by the current tendency of incorporating two or more barbiturates, with

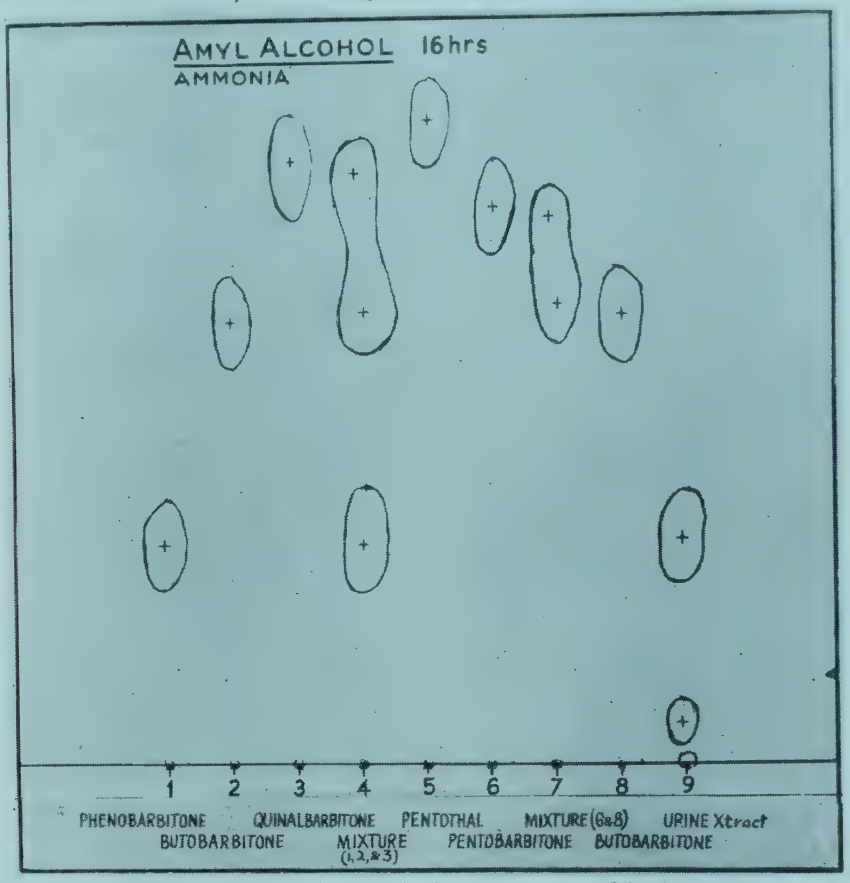


Figure 6. The chromatogram shows (*a*) the separation of barbiturates present in compound commercial preparations, and (*b*) a urine extract containing phenobarbitone and a metabolite.

different rates and times of hypnotic action, into a single tablet. Prior to the advent of paper chromatography, all barbiturates had to be estimated as 'barbitone' even when strong evidence pointed to the ingestion of a particular barbiturate. However, it is now possible without too much difficulty to determine which drug or drugs have been ingested, from the position of the unchanged drug on the chromatogram and to confirm its identity from the positions of its metabolites.

As has been stated previously, the detection of sugar in 50 μl. urine is abnormal, and there is nothing whatsoever to be gained by determination of the quantity present. Indeed, wherever the diagnosis depends simply on the appearance of a compound (sugar, amino-acid or indole) when an equal volume of normal urine should show none, then qualitative analysis is sufficient and it is pointless to quantitate the result. However, where the abnormality is reflected, not in the appearance or disappearance of a compound or compounds but in quantitative variations, then obviously the analysis must be quantitative or nothing is accomplished. Examples of this latter case are the steroids and the iodine compounds, and we shall consider the steroids in a little more detail. The oestrogen and pregnanediol content of children's urine is vanishingly small but, with the onset of puberty in the female, the cyclical excretion of both begins. Again in adrenal hypoplasia the excretion of steroids will be low when compared to normal while in hyperplasia of the gland the excretion is increased. Various forms of adrenal hyperplasia result in the excretion of increased amounts of different steroids; the adrenogenital syndrome results in an excess of urinary pregnanetriol whereas hyperplasia causing Cushing's syndrome shows excess cortisol excretion. Originally a total steroid figure was determined, e.g. total cortical steroids, but it is now generally realized that this figure is much less meaningful than was originally thought.

We may turn now to those families of compounds which require two-way chromatography for their investigation. These include the amino-acids, indoles and phenolic acids. However it should be noted that one-way chromatography is sufficient to control or examine the therapeutic measures adopted in many cases.

Normal urine shows about six amino-acid spots on a two-way chromatogram when 100 μl. of de-salted urine is run, and the pattern obtained serves as a guide to the identity of abnormal spots when they appear. Although the normal amino-acids vary in concentration between individuals, no condition is known where the abnormality is reflected in a large variation in the concentration of one of these normal compounds. Two general types of variation in pattern are observed; those where one or two particular abnormal spots appear and those where there is a generalized excretion and many spots are present. In the former case are included cystinuria, when cystine and arginine or lysine or both are found in addition to the normal amino-acids, and phenylketonuria when the only usual spot is due to phenylalanine; often the extra spot is present in such large amounts that the volume of urine used for chromatography

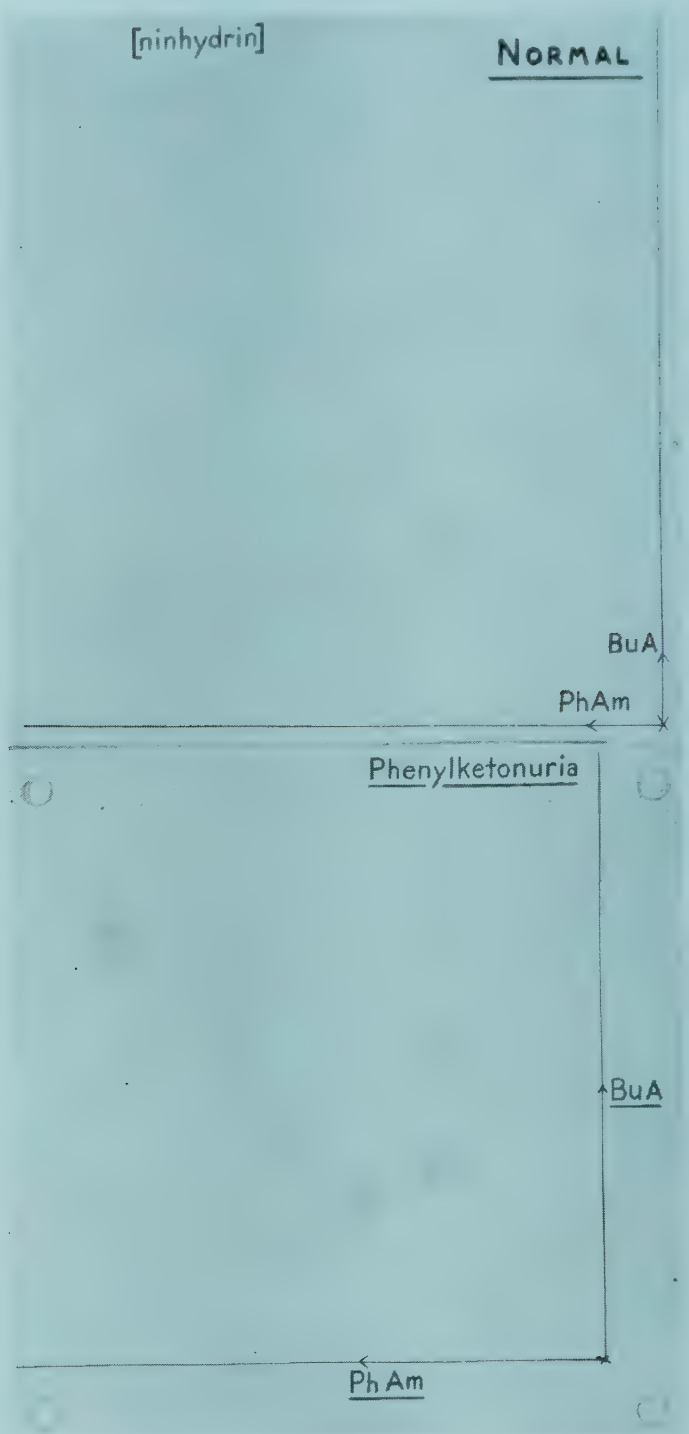


Figure 7. Chromatograms of a normal urine and of various abnormal specimens illustrating the appearance of one or many abnormal amino-acid spots.

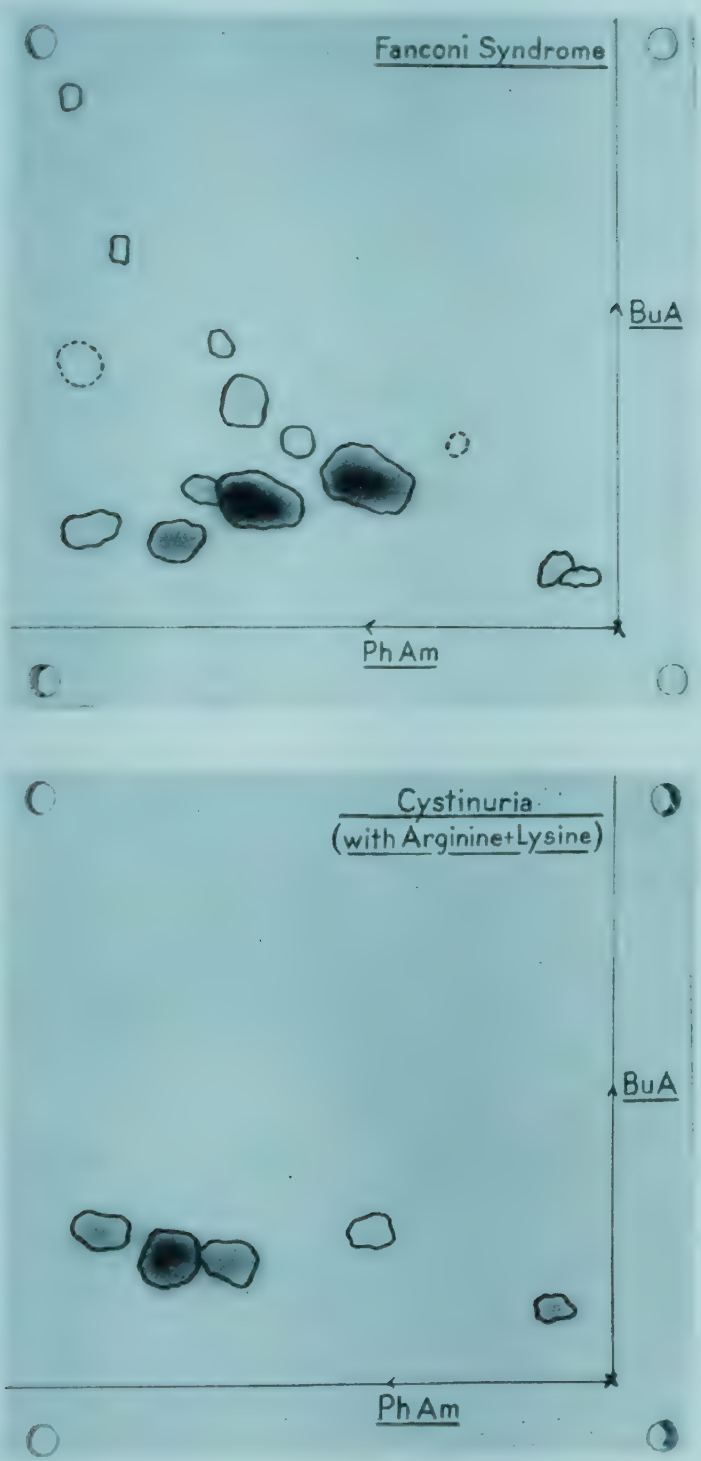

Fanconi Syndrome
BuA
Ph Am
Cystinuria
(with Arginine+Lysine)
BuA
Ph Am

Figure 7

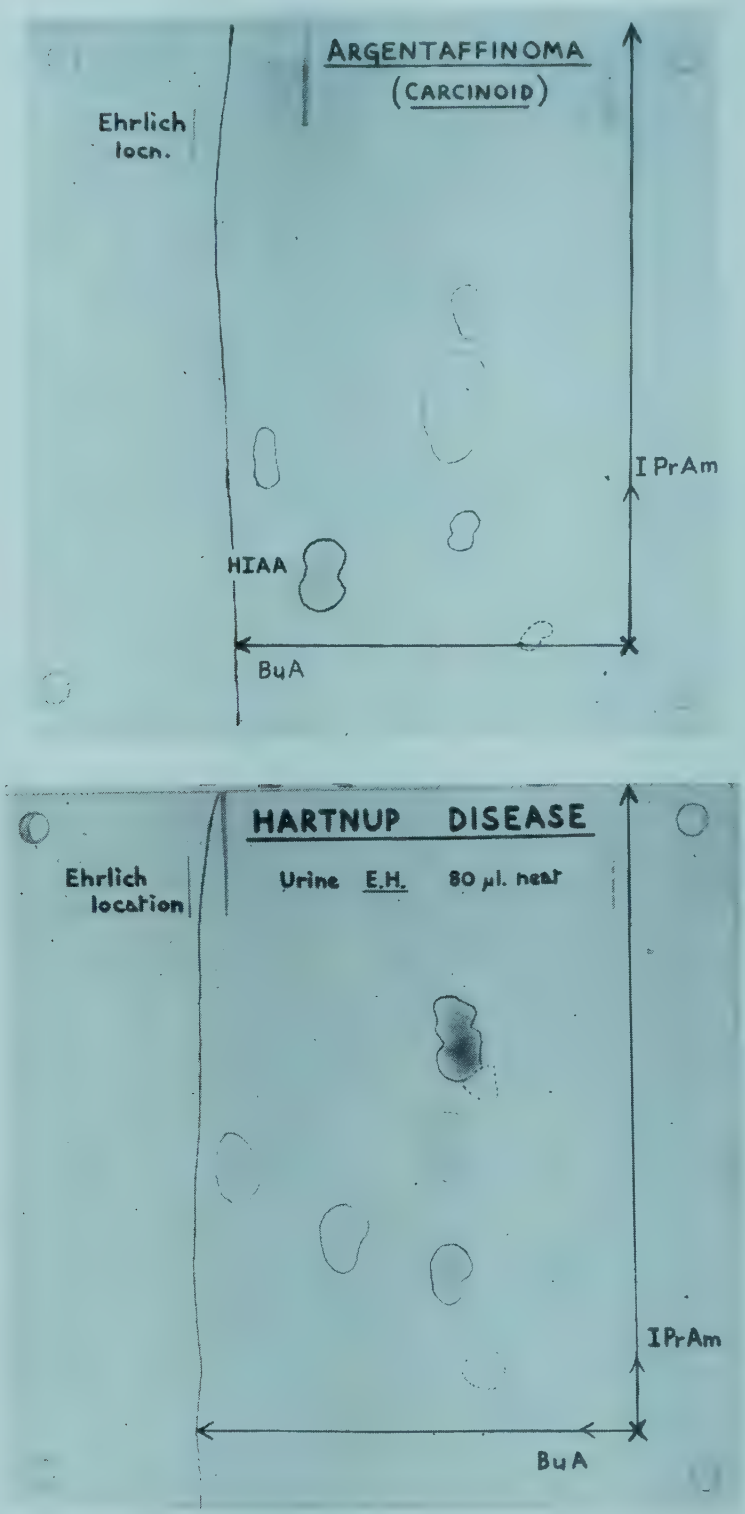

ARGENTAFFINOMA
(CARCINOID)
Ehrlich
loch.
IPrAm
HIAA
BuA
HARTNUP DISEASE
Urine E.H. 80 µl. neat
Ehrlich
location
IPrAm
BuA

Figure 8

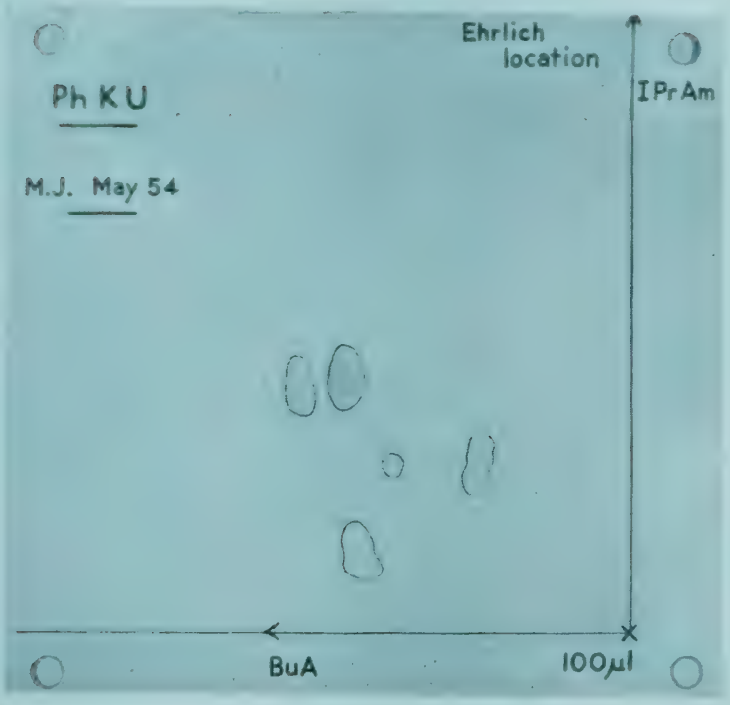


Figure 8. Chromatograms of various abnormal indolurias illustrating diagnostic patterns. The small number of different spots in quite different patterns makes interpretation relatively simple.

can be cut down with the result that the normals are barely visible on the chromatogram. In the latter case are included Hartnup disease, Wilson's disease, Fanconi syndrome, various necroses and vitamin deficiencies. The generalized amino-aciduria may be associated with liver disease or renal tubular defects or both and these may be genetic in origin or acquired.

Urinary indoles are perhaps the simplest family of compounds to chromatograph as the 'neat' urine can be applied directly to the chromatogram without the need for preliminary de-salting. On chromatographing some 100 µl. of urine on a two-way run, only urea, indoxyl sulphate (urinary indican) and tryptophan are present in sufficient amounts to show up with the usual reagents. To date only three conditions are known to yield characteristic abnormal patterns, namely argentaffinoma, Hartnup disease and phenylketonuria, although abnormal indole excretion has been implicated in other conditions. Chromatographic diagnosis of argentaffinoma is based on the appearance of an enormous spot due to

hydroxyindolylacetic acid although more advanced cases show other metabolites. In phenylketonuria spots due to indolyllactic acid and large amounts of indolylacetic acid are also present. This indole pattern is less valuable than the appearance of phenylalanine for identification purposes and the latter is a more suitable pointer during the therapeutic control. Hartnup disease reflects itself in the appearance of a triad of indolylacetic acid, indolylacetylglutamine (the detoxification product of indolylacetic acid) and a greatly raised tryptophan.

In contrast to the urinary indoles, the phenolic acids are amongst the most difficult to interpret. These compounds are present in concentrations such that the equivalent of 1 ml. of whole urine must be applied to the origin of a two-way chromatogram. The normal pattern shows some twenty to thirty spots including a number which have not yet been identified. Further, the variation is such that the normal pattern can only be deduced from an examination of the many slightly different patterns obtained from known normal individuals. Much of this variation is dietary in origin; for example, aspirin leads to the appearance of a salicylic acid spot, vegetables and coffee have been claimed to lead to the appearance of other spots. In spite of this it is possible to associate certain variations in the qualitative and quantitative patterns with particular disorders or derangements of the metabolism of aromatic compounds. Thus in phenylketonuria a number of ortho- and paraphenolic acids increase in intensity whereas spots due to other acids diminish. Ascorbic acid is concerned in the metabolism of aromatic compounds and, therefore, where a deficiency of this vitamin occurs or an unsatisfied extra requirement manifests itself, derangements of aromatic metabolism occur and are reflected in the chromatographic pattern. Hence, scurvy, thyrotoxicosis and steatorrhoea, to name but a few conditions, result in abnormal patterns.

Having determined the nature of the disease, it may be possible to prescribe some therapeutic measure which results in a return to health and consequently to the normal pattern. In galactosemia one of the enzymes is absent with the result that this compound cannot be converted to glucose in the normal manner. This results in a toxic build-up of galactose, first in the blood and then as an overflow into the urine. The galactose, of course, is derived from maternal-milk lactose and later from cow's milk. Therapy is therefore quite simple, namely the removal of milk and milk products from the diet. The efficacy of the treatment can be ascertained by periodic chromatography when galactose should be quite absent. Similarly, the success and control of phenylalanine

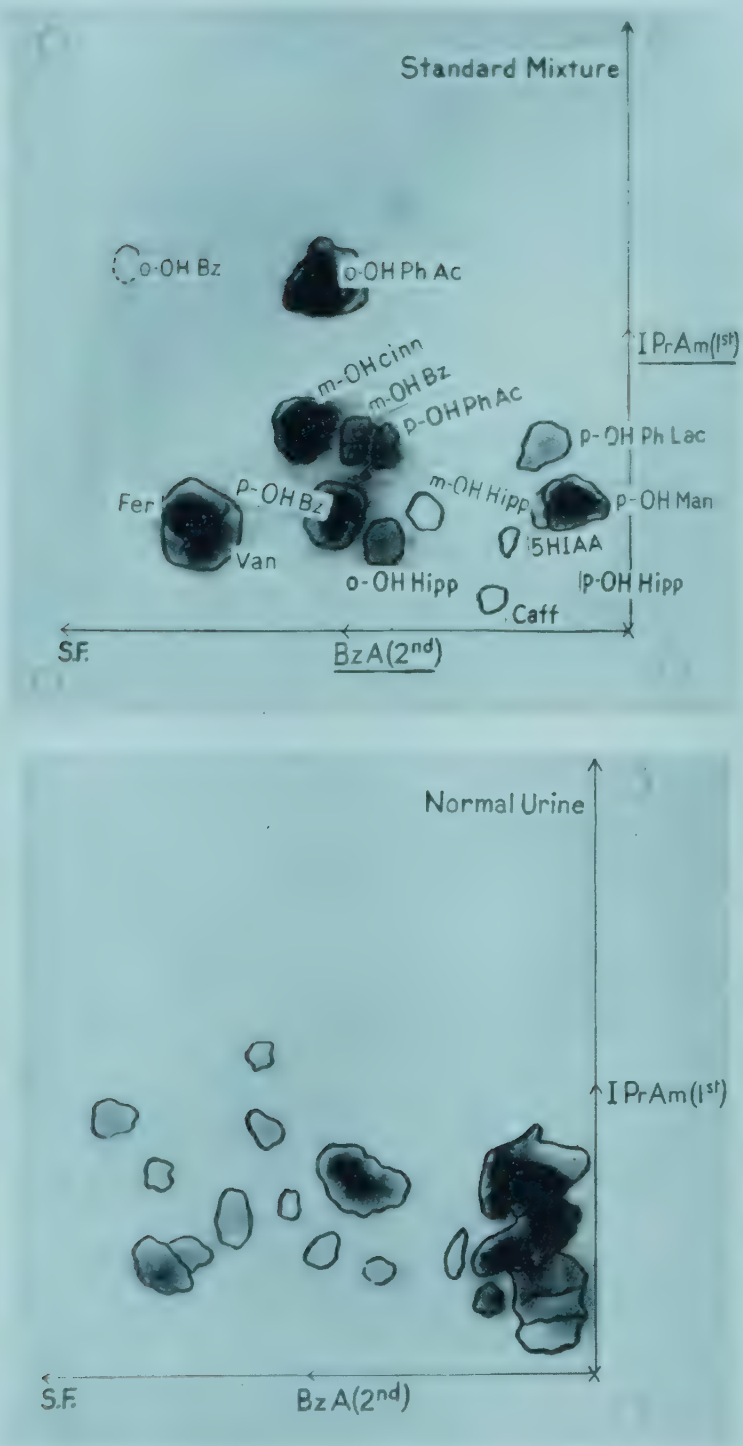


Figure 9. Chromatogram of a standard synthetic mixture of phenolic acids, and the pattern obtained from a normal urine. The large number of overlapping spots renders interpretation comparatively difficult.

low diets in the treatment of phenylketonuria can be gauged by periodic one-way blood chromatography. If the amino-acid is present then the phenylalanine content of the diet must be reduced. Conversely if the chromatogram shows none to be present then the dietary concentration can, if thought necessary, be increased to provide basal requirements.

These examples illustrate a further advantage of the chromatographic method. Derangements in the metabolism of galactose and phenylalanine produce irreversible brain damage and so it is of the utmost urgency to determine such conditions as early as possible after birth. The total time required for such determinations is less than twenty-four hours and, at the same time, it enables a tolerance test to be eliminated.

Attention may here be drawn to the fact that the diet can influence the urinary components. For example, false positives may be obtained after the ingestion of berry fruits which result in the appearance of the sugars, xylose and arabinose, in the urine. The ingestion of bananas can lead to a false diagnosis of argentaffinoma due to the large amounts of hydroxy-tryptamine in that fruit. Dates may lead to the appearance of a 5-hydroxy pipecolinic acid on amino-acid chromatograms. In all cases, however, dietary restrictions lead to the disappearance of these spots.

In all the above cases we have been concerned to show the presence of one or more abnormal substances on the chromatogram for diagnostic or therapeutic purposes; thus we have taken the presence of galactose to indicate that the patient was galactosemic. The use of paper chromatography, however, is of at least equal importance in the study of both normal and abnormal metabolism. To illustrate this we will take the example of phenylalanine metabolism as it is fairly well known and, also, this metabolism may go awry at a number of different stages along the pathway.

In the normal, phenylalanine in excess of body requirements is metabolized through tyrosine, *p*-hydroxyphenylpyruvic acid, homogentisic acid to carbon dioxide and water. Information of this type is determined in the following way. Either a large dose of the particular compound or a minute dose of the radio-actively labelled compound is given, and chromatograms of the urine and blood are prepared. Increased amounts of metabolites, or labelled metabolites, but not of other substances, will appear on the chromatogram. Thus if labelled *p*-hydroxyphenylpyruvic acid is given, then labelled homogentisic acid appears but not tyrosine or phenylalanine and so the metabolic sequence is established. Of course, if the spot on the chromatogram cannot be identified by means of R_f and colour reactions as a known substance, it must be worked up in larger

amounts and purified by elution as previously described so that its structure can be determined by standard chemical methods. Abnormalities are known in which each of the four stages of phenylalanine metabolism shown are blocked. In phenylketonuria, for example, the para-hydroxylating enzyme which converts this amino-acid to tyrosine is absent (position A in the figure) and therefore the phenylalanine concentration builds up in the body. Tyrosine metabolism is normal as shown by the fact that even large amounts of ingested tyrosine are metabolized in the normal manner. When any substance, even a normal one, builds up in the body it becomes toxic and the body responds by attempting to

normal pathway	Phenyl alanine	A →	Tyrosine	B →	*p*-OH phenyl pyruvic acid	C → homogen- tisic acid	D → $CO_2 + H_2O$
abnormal pathways	Phenyl- lactic, acetic acids and *o*-OH analogues				*p*-OH phenyl- lactic and acetic acids		
abnormal conditions	phenyl- ketonuria				tyrosinosis scurvy steatorrhea	alkaptonuria ochronosis	

Figure 10. Normal and abnormal pathways in phenylalanine metabolism. Phenylalanine is normally metabolized through tyrosine, *p*-hydroxyphenylpyruvic acid, homogentisic acid and aceto-acetate to carbon dioxide and water. Metabolic blocks can occur at stages A, C and D. In the two former cases, the compounds before the block are metabolized through abnormal or unusual pathways whereas in the latter case the compound concerned is not metabolized further but excreted as such.

eliminate it via other abnormal pathways. As a result of this, one or more unusual spots may appear on the chromatogram and these can be identified using the methods described above for normal metabolites. In phenylketonuria, one finds both the original phenylalanine plus a large number of non-hydroxylated aromatic acids as well as a number of ortho-hydroxylated aromatic acids although these are present in smaller amounts. If we consider another block in the same sequence (position C in the figure) we find a build up of *p*-hydroxyphenylpyruvic acid, with the following consequences. First, abnormal metabolites of this compound, namely

p-hydroxyphenyllactic and *p*-hydroxyphenylacetic acids appear and, second, none of the phenylalanine abnormal metabolites appear as the metabolism of this latter compound is normal.

Certainly the most striking recent application of paper chromatography has been in the elucidation of the partial or complete structure of peptides and proteins. By means of the procedure devised by Edman, it is now possible to split off one amino-acid residue at a time from a peptide chain; this is illustrated in the figure:

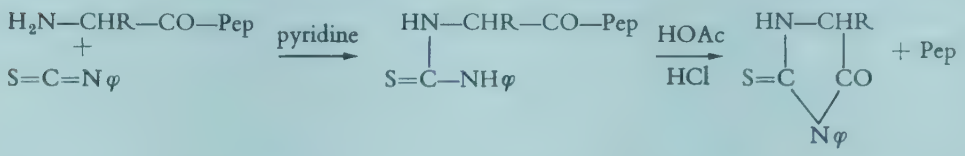


Figure 11. Edman stepwise degradation of peptides. The peptide is reacted with phenyl*iso*thiocyanate (Q—N=C=S) in an atmosphere of pyridine. The product is then cyclized by placing in an atmosphere of glacial acetic acid-hydrochloric acid when the end amino-acid is split off as the phenylthiohydantoin derivative. The remaining peptide, reduced in length by one amino-acid, is then separated from the hydantoin and resubmitted to the whole procedure when a second amino-acid is split off, and so on. The liberated hydantoin is determined by paper chromatography.

For some as yet unexplained reason, this reaction ceases sometimes after only two or three residues have been removed, at other times when as many as ten to twelve residues have been split off one at a time. Use is then made of enzymes which preferentially split particular bonds in the chain, in order to obtain simpler peptides which can be separated either chromatographically or electrophoretically and then each separately resubmitted to the Edman procedure.

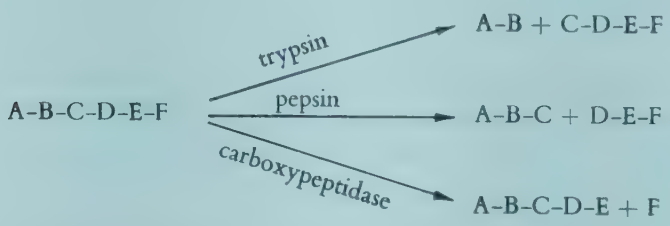


Figure 12. Illustration of the enzymic splitting of a peptide, A–B–C–D–E–F. Enzymes act by splitting peptide bonds associated with particular amino-acids. Thus, in the illustration, the enzyme carboxypeptidase splits off the amino-acid at the end of the chain containing the free carboxyl group; trypsin attacks bonds between the amino end of arginine or lysine and the carboxyl end of any other amino-acid, etc.

In this way, the structures of the various peptides, A–B, C–D–E–F, etc. can be separately deduced and the complete structure elucidated. Using this and similar methods, the complete amino-acid sequences in corticotrophin, α and β MSH, oxytocin, vasopressin and insulin were deduced. Similarly, partial sequences of higher molecular weight proteins and enzymes are being deduced. Furthermore, as the Edman procedure is so gentle, it is sometimes possible to carry out bioassay of the degraded molecule in order to determine whether the residue is active or whether there is an 'active core' concerned with the biological properties of the individual molecule. Thus it is possible to remove nineteen amino-acid residues from the papain molecule without loss of enzymic activity.

References to this whole subject may be found in *Chromatographic and Electrophoretic Techniques*, Heinemann Medical Books, London; and Interscience, New York, 1960, in which all facets of the practice of chromatography and its clinical and biochemical applications will be found. Most of the figures in this paper are taken from the above book.

THE ULTRACENTRIFUGE

PALEY JOHNSON

SEDIMENTATION IN GENERAL

If a coarse clay suspension is homogenized by shaking and then allowed to stand, the settling under gravity of the clay particles can be observed by following the motion of the well-defined 'boundary' (between clay suspension below and clear solvent above) which forms and gradually moves downwards. Such observations have, for many years, been the basis of methods used for determining the size of the colloidal particles, and they are applicable to a range of particle sizes, most of which is within microscopic visibility. However, as the particle size decreases it is found that the motion of the boundary becomes slower, and the boundary itself becomes less sharp; finally the boundary becomes so broad that its motion cannot be followed by simple means. The reason for this behaviour is that, with decreasing particle size, the movements of Brownian motion become considerable and Brownian displacements eventually become comparable or even greater than those due to gravity. This must lead to ill-defined boundaries and it is clear that for the most favourable application of the rate of sedimentation under gravity as the means of estimating particle size, the displacement under gravity must be considerably in excess of that due to Brownian motion in the same time-interval. The range of the method varies from system to system with the density difference between particle and solvent—the larger the density difference, the smaller the limiting particle size. It is worth noting that where the particle density is lower than that of the solvent, then sedimentation occurs upwards (as in the separation of cream from milk) but this case is not fundamentally different from the more usual sedimentation downwards. In the remainder of this paper, unless otherwise stated, it will be assumed that particle density is the greater and for such systems the limiting equivalent diameter for successful use of the method is usually taken to be about 10^{-4} cm.

To extend the method to smaller particle sizes, it is necessary in some way to increase the displacements under sedimentation relative to those

of Brownian motion. Whilst a lowering of liquid density has certain limited applications, the replacement of the gravitational by a high centrifugal field offers great possibilities. The ultracentrifuge is merely a means of achieving such fields, at the same time allowing detailed observation of sedimentation, which must occur without any convectional or other disturbances. With the high centrifugal fields now available, the effective limiting equivalent diameter can be as low as 10 Å (10^{-7} cm.) so that a thousand-fold extension of the method is achieved.

The Ultracentrifuge

In all ultracentrifuges, a cylindrical steel or duralumin rotor (Fig. 1) holds a cell, containing a sector-shaped cavity which is carefully adjusted to be radial. It is thus ensured that the path of sedimenting material at the radial walls is parallel with the walls—misalignment or the use of other shapes of cavity are liable to cause convectional disturbances. The rotor is spun, usually about a vertical axis and usually through a flexible steel shaft (Fig. 2) which makes the balancing of the rotor much less critical than would otherwise be necessary. The shaft is itself spun by an air turbine or some type of electrical motor. In the former the air turbine is lifted slightly from its rest position by an air 'cushion', thus to a very

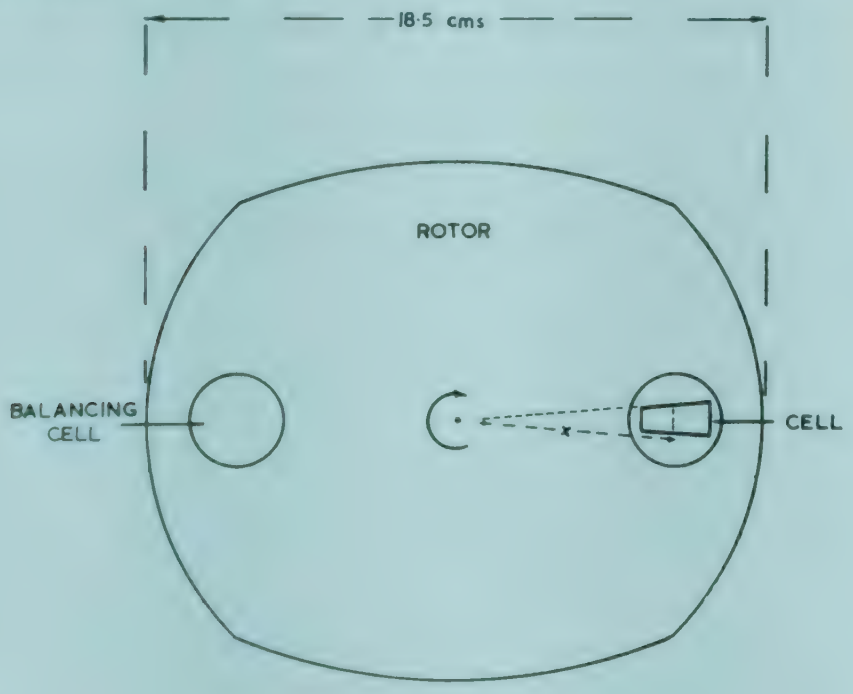


Figure 1. The ultracentrifuge rotor (plan).

large extent eliminating high-speed frictional trouble, and is driven to high speed by air pressure. In the most frequently used (Spinco) ultracentrifuge of the day, the shaft is driven by a cooled high-speed electric motor (12,000 r.p.m.) geared up by a factor of 5. In both cases, the rotation must be smooth and without vibration. The maximum routinely achieved rotational speed for a 18·5 cm. diameter rotor is 60,000 r.p.m., which corresponds roughly to 250,000 g in the cell. At higher speeds, so great is the centrifugal field near the periphery of the rotor, that there is a definite risk of the rotor disintegrating. To achieve high speeds, the

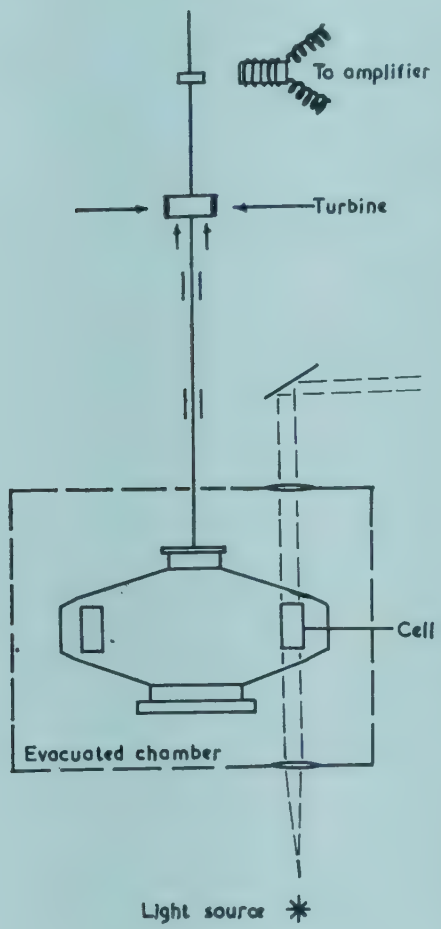


Figure 2. The air-driven ultracentrifuge (diagrammatic).

rotor chamber must be continuously evacuated (to less than 10^{-3} mm. Hg) or excessive air friction and heating of the rotor would be experienced. Oil glands around the driving shaft effectively seal the chamber from the drive system above. Rotational speed is measured, as indicated in Fig. 2,

by the use of a small electromagnetic generator on the drive shaft, the frequency of the alternating signal produced being measured to a high degree of accuracy by standard methods. Alternatively, a revolution counter connected through a suitable gearing system to the main drive, may be used over a convenient time-interval to give the mean speed for the interval.

The temperature of the rotor is usually measured to $\pm 0{\cdot}1^{\circ}$ C. by a thermistor placed very close to the rotor surface (an elaborate calibration being then necessary to ensure the correctness of the readings) or inside its base, electrical contact then being obtained by a tungsten needle dipping into a pool of mercury. It is essential, if undisturbed sedimentation is to occur, that the temperature of the cell be steady or change only very slowly, and that the rotor be free from excessive temperature gradients. The cell (Fig. 3) contains the sector piece (s), a cylindrical component

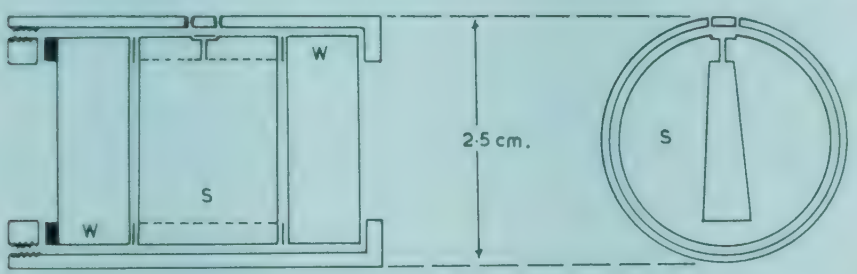


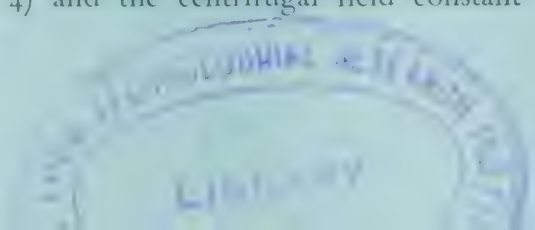
Figure 3. The ultracentrifuge cell.

in anodized duralumin or plastic taking up to 1 ml. of test liquid, provided with strong good quality quartz windows (w) at each end so that the contents of the cell may be observed optically during sedimentation. Special care in providing solvent-resistant sealing gaskets between the various components is essential for the prevention of leakage at high centrifugal fields. Fig. 2 shows the general disposition of components of the optical systems which are used to follow sedimentation.

Summarizing, the ultracentrifuge provides a means whereby a small quantity of test liquid is subjected to a high centrifugal field at a constant temperature and in which sedimentation, uncomplicated by vibrational or convectional disturbances, may be observed.

SEDIMENTATION IN THE ULTRACENTRIFUGE CELL

In understanding the main features of sedimentation, we can imagine the cell to be rectangular (Fig. 4) and the centrifugal field constant



throughout the cell (strictly it is proportional to the distance x from the axis of rotation). If the solution of macromolecules is dilute and monodisperse, then all macromolecules will move in the direction of the field with constant velocity, v. Assuming (for the moment) diffusion to be negligible, then after time t, a 'boundary' will occur at a position distant vt from the meniscus at which solvent meets solution (Fig. 4(a, i)). With increasing time, the boundary moves down the cell with velocity v and clearly, if this velocity can be determined, it provides the velocity of the macromolecule under the conditions of experiment. Fig. 4(a, ii) gives the distribution of concentration (c) in the cell at a given time and Fig. 4(a, iii) the slope (dc/dx) as a function of distance (x) along the cell. For a sym-

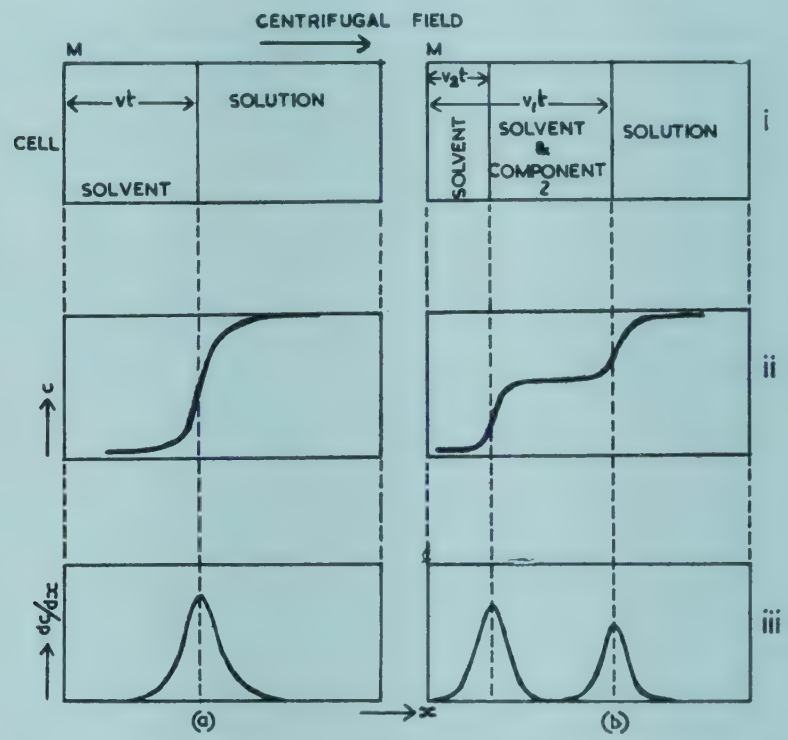


Figure 4. Sedimentation of (a) monodisperse solute; (b) mixture of two monodisperse solutes.

metrical peak (as in 4(a, iii)) it can be shown that the maximum of the peak corresponds with the point of inflexion of the c v x curve and with the position of the boundary in Fig. 4(a, i). The optical systems used are designed to give information of the type contained in Figs. 4(a, ii and a, iii). Chiefly used are the diagonal Schlieren optical systems, based upon refractive index differences, which give patterns of the latter type directly

on a viewing screen or photographic plate. A precise location of the position of the boundary is thus provided. Fig. 5(*a*) shows such patterns for a haemoglobin solution in which the Schlieren peaks are superimposed upon the absorption pattern of the strongly coloured protein solution. Detailed examination of the absorption could be used to give the concentration throughout the cell but the Schlieren patterns are more readily obtained and informative. In Fig. 5(*a*) it can be seen that in addition to moving from left to right across the cell, the peak spreads with time and by recalling the derivation of dc/dx v. x curves (see Fig. 4(*a*)), it is seen that broader peaks are associated with a less steep variation of c with x. This result is just what might be expected for later times from diffusion,

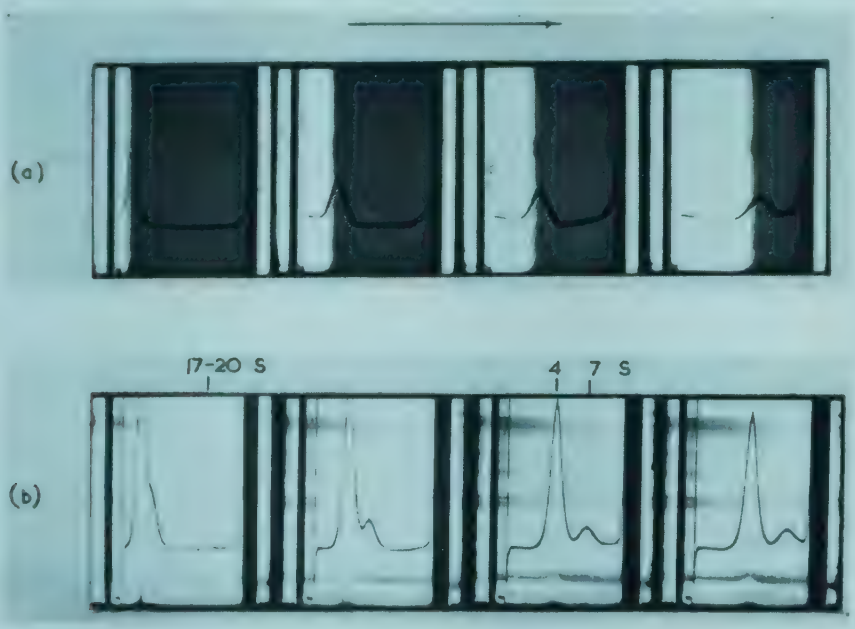


Figure 5. Diagonal Schlieren patterns at 59,780 r.p.m. for
a. Sheep haemoglobin in borax (0·2 M), NaCl (0·1 M) buffer at pH 9·1. Protein conc., 0·38 gm./100 ml.
b. Normal human serum diluted to 5 volumes in 0·06 M, barbiturate buffer at pH 8·6.

which tends to smooth out concentration differences where they occur. Thus we have diffusion superimposed upon the motion of the peak arising from sedimentation. This is usually the simplest possible system.

If now we have two types of macromolecule in solution which move independently of one another with velocities v_1 and v_2 under the conditions of the experiment, we have two boundaries forming whose separation $((v_1-v_2)t)$ increases with time and for which curves of c and dc/dx

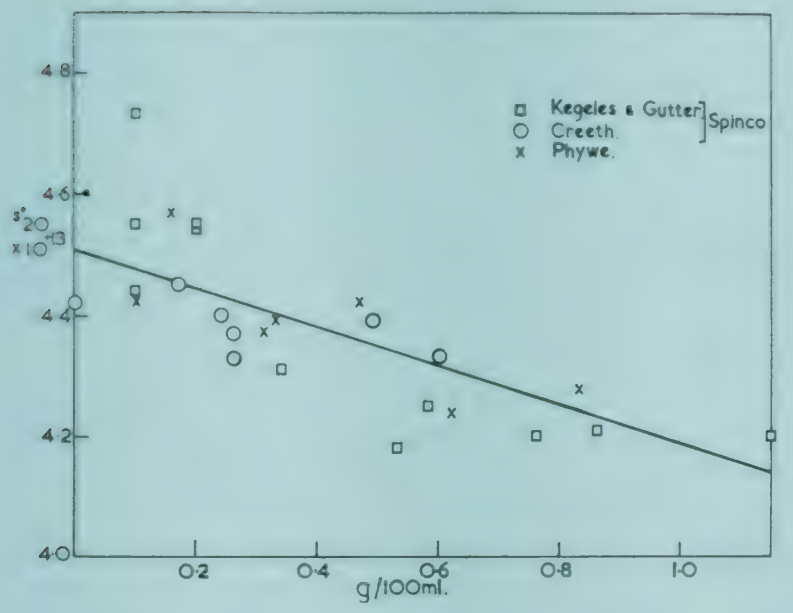
against x are given in Fig. 4(b, ii and iii). Diagonal Schlieren patterns for such a system are given in Fig. 5(b) for a normal human serum. The increasing separation with time of the two main peaks can be followed quantitatively to give information on the motion of the molecules responsible and it can be shown that the areas of the peaks are related (often proportional) to their concentrations. Thus the Schlieren methods provide information directly on the positions of the moving boundaries and on the concentrations of the species responsible for them. Clearly resolution of the peaks depends upon the occurrence of a sufficient velocity difference and in any particular experiment there is a limit to the number of components which can be resolved. If a large number of different molecular species (i.e. a polydisperse system) occur with sedimentation velocities too close for resolution, then the initially sharp boundary (which always occurs if the sedimentation of the solute relative to solvent is appreciable), spreads faster than is consistent with diffusion and to an extent which is determined by the range of velocities represented.

SEDIMENTATION AND MOLECULAR SIZE

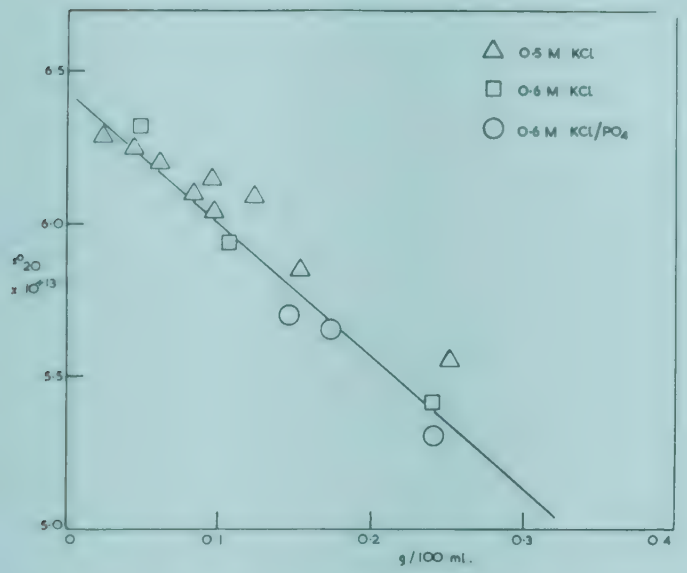
To allow comparisons between experiments under different conditions, the sedimentation velocity, determined from photographic records of the motion of a boundary, is divided by the centrifugal field ($\omega^2 x$, where ω is angular velocity of the rotor) to give the sedimentation velocity per unit field, or sedimentation constant, s. Such a constant characterizes the steady motion of the macromolecule in the centrifugal field and the frictional forces associated with such motion per molecule can be equated to the centrifugal force experienced by the molecule in the solvent. It is implicitly assumed here that solute molecules do not interfere with one another and to ensure that experimental results conform with this condition, it is necessary to work at low concentration (c) of solute ($\ll$ 1 g/100 ml.) and advisable to extrapolate to zero solute concentration (as in Fig. 6) to give $[s]_{c=0}$. It can thus be shown that

$$M = \frac{RT[s]_{c=0}}{[D]_{c=0}(1-\bar{v}\rho)}$$

where M is molecular weight, R is the gas constant, T is absolute temperature, ρ is the solvent density and $\bar{v}$ the partial specific volume of solute (or reciprocal of particle density). The term $(1-\bar{v}\rho)$ accounts for the buoyancy of the solvent which clearly is involved. $[D]_{c=0}$ is the diffusion



(a)



(b)

Figure 6. Plots of sedimentation constant corrected to water at 20° C. (s°_{20}) against protein concentration for (*a*) bovine serum albumin, (*b*) rabbit myosin.

coefficient, determined under (or corrected to) the same solvent conditions as were used for the sedimentation, and, like the sedimentation constant, it requires extrapolation to zero solute concentration. In principle it may

be determined from the spreading of the boundary during sedimentation but is more usually and more accurately determined by a separate experiment in which boundary spreading occurs in a special stationary diffusion cell. In calculating molecular weights from sedimentation data, the above equation is normally utilized and it yields molecular weights which have not involved any assumption regarding molecular shape and which approximate very closely to the unsolvated condition (even though in solution the molecule may be solvated). Table I contains representative sedimentation and diffusion data with derived molecular weights for the better known proteins.

Proteins are chosen for illustrative purposes not only for their obvious biological importance, but because their physical properties (including molecular weight) possess a high degree of constancy; with time, from individual to individual, and frequently amongst different species. It is not necessarily suggested that any particular protein system is monodisperse (i.e. comprised of identical molecules) but if, for any particular property, a distribution around a mean value occurs, the distribution is usually sharp and constant.

Insulin monomer is regarded as one of the lowest molecular weight proteins but protein molecular weights extend well into the millions. Amongst other types of macromolecule, less detailed information is available but nucleic acids and nucleoproteins are also known to have very high and frequently variable molecular weights.

It may be enquired—do not molecular shape and solvent-binding capacity affect the processes of sedimentation and diffusion? Such properties do have an important influence and mere calculation of molecular weights does not utilize sedimentation and diffusion data fully. It is readily possible to calculate, for a particular system, a *frictional ratio* which gives a measure of the frictional resistance offered by the solvent to the molecule in question as compared with that for a hypothetical unsolvated spherical molecule of the same molecular weight. For a real molecule which is spherical and unsolvated such a ratio has the value unity; but real systems almost invariably give a ratio which is significantly greater than unity (Table I, column 5). Further progress here is speculative. If we can treat the molecule in terms of some simple geometrical shape, particularly an elongated ellipsoid of rotation, and can estimate the solvation (i.e. the firm binding of solvent molecules by the macromolecule), then it is possible to obtain information on the particle shape. Column 6 of Table I contains the axial ratio for the equivalent elongated ellipsoid of rotation assuming that 1 gm. of dry protein firmly binds 0·3 gm. of water. Whilst

some of the lower molecular weight proteins behave as hydrated spheres, most must to a greater or less extent be asymmetric. The extreme asymmetry of fibrinogen, with its specialized biological function, is of great interest.

TABLE I. *Molecular data from sedimentation and diffusion measurements*

	$s^{\circ}_{20} \times 10^{13}$	$D^{\circ}_{20} \times 10^{7}$	$M \times 10^{-3}$	Frictional ratio	Axial ratio*
Insulin (monomer)	1·6	15	12·0	(0·95)	1
Insulin (trimer)	3·5	8·2	41·0	1·13	1
Ribonuclease	1·85	13·6	12·7	1·04	1
Lactalbumin (cow)	1·9	10·6	17·4	1·17	2·0
β-Lactoglobulin	3·12	7·30	41·5	1·26	3·2
Ovalbumin	3·55	7·76	44·0	1·16	1·9
Serum albumin (cow)	4·45	6·05	67·5	1·31	3·6
Carboxy haemoglobin (man)	4·46	6·9	63·0	1·16	1·9
Diphtheria toxin	4·6	6·0	74·0	1·22	2·6
γ-Globulin (man)	7·1	3·84	176	1·49	6·5
Fibrinogen (cow)	7·9	2·02	330	2·34	23
Myosin (rabbit)	6·43	1·05	540	3·76	67
Legumin (peas)	13·7	3·02	410	1·43	5·5

* Assuming the molecule to be an elongated ellipsoid of rotation with 30 per cent hydration.

Information of the above type is obtained usually on highly purified systems, so that possible interactions from other types of macromolecules are avoided. However, the ultracentrifuge is probably used to a greater extent in investigating systems, particularly biological fluids, which contain a variety of different macromolecules. Less detailed information on particular components is thus acquired but much new knowledge becomes possible, e.g. the composition of the whole system in terms of the differently sedimenting component, changes in it occurring over a variety of conditions, and the nature of the interactions which may occur between different components. In the remainder of this article such possibilities will be illustrated. Whilst advocating the considerable possibilities of the ultracentrifuge, it should be realized that materials of similar sedimentation properties will not be so resolved and great advantages accrue when other methods, based upon different physical principles, are used in parallel. Particularly is this true of electrophoresis, in which the motion of dissolved molecules is determined by electrical properties; electrophoretic patterns may thus bear little or no superficial resemblance to those from the ultracentrifuge.

SERUM SEDIMENTATION PATTERNS

Fig. 5(*b*) contains sedimentation diagrams for a normal human serum. Three differently sedimenting components can be readily detected. The most abundant (*ca.* 80 per cent of total area), with a sedimentation constant of about 4 S (1 S = Svedberg unit = 1×10^{-13} c.g.s. units) contains largely albumin, β-globulin and at least some of the α-globulin. Clearly this is a case where several components with similar sedimentation properties have distinctly different electrical properties and are readily resolved in an electric field. The next most abundant (*ca.* 16 per cent of total area), of sedimentation constant about 7 S, contains largely γ-globulin. The trace of rapidly sedimenting material (*ca.* 4 per cent of total area), often termed macroglobulin, has a sedimentation constant of 17–20 S and its relation to electrophoretic components is not completely established—in some cases it appears to be an α-globulin, in others it has γ-globulin properties. In this case the ultracentrifuge provides the greater resolving power, for clearly within a single electrophorectic component we may have molecules of different sedimentation behaviour. It seems clear that, amongst normal sera, some variation in the amounts of the components and in their specific properties must be allowed, but gross alterations of either type and, in addition, the appearance of different (normally absent) components is indicative of biological disturbance. During the last few years, in collaboration with the Pathology Department of Addenbrooke's Hospital, Cambridge, we have been examining different types of abnormal sera with the idea, ultimately, of classifying the resulting serum patterns and possibly relating to clinical condition. This goal has not, by any means, been reached but a bewildering variety of possible serum patterns has been recorded. Fig. 7(*a*) demonstrates that the pattern may vary from almost single 4 S component type (little 7 S) through all possible intermediate stages to almost single 7 S component (little 4 S component). Such a variation involves a considerable change in osmotic pressure per unit of protein concentration, but total protein concentration is frequently abnormal in such cases and must be considered.

Not only do we observe variations in the albumin/globulin ratio but minor components may appear to a major extent or even completely new and normally absent components (Fig. 7(*b*)). Thus in Fig. 7(*b*, i) the proportion of the 18–20 S component is some ten times as great as normally, though it cannot be concluded that the component in question is identical with that normally occurring. Identity of sedimentation behaviour is not a proof of complete identity. Parallel electrophoretic examination indi-

cates the occurrence of a considerable proportion of a more mobile γ-globulin. The clinical picture associated with the pattern of Fig. 7(*b*, i) is somewhat indecisive and Waldenstrom (1944), regarding the sedi-

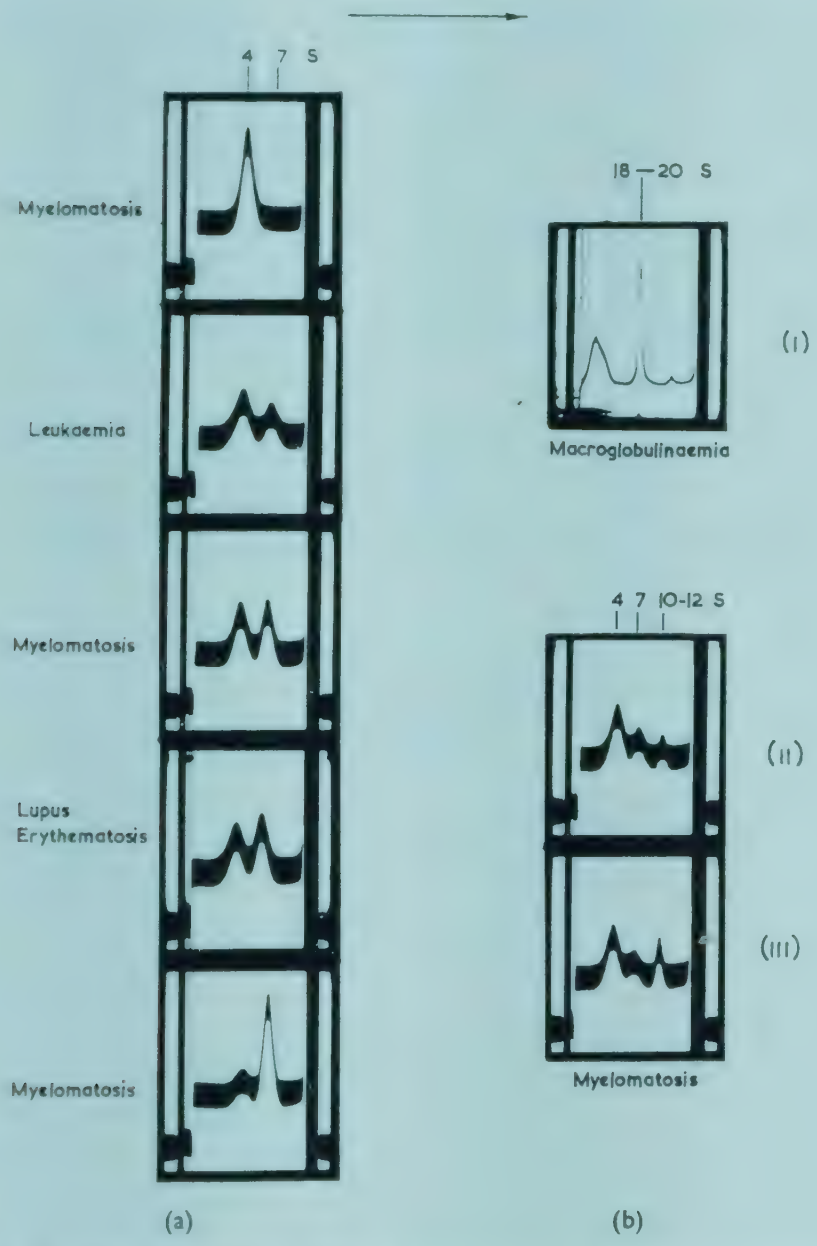


Figure 7. Patterns at comparable stages of sedimentation for various pathological sera
Dilution to 5 volumes in 0·06 M barbiturate buffer at pH 8·6, 59,780 r.p.m.
a. Showing variation in 'Albumin/Globulin' ratio.
b. Showing occurrence of unusual sedimenting components.

mentation pattern as the most definite feature, refers to the condition as macroglobulinaemia, in view of the undoubted high molecular weight of the increased 18 S component. Fig. 7(*b*, ii and iii) show the appearance to different extents of a component of sedimentation constant about 10–12 S.

The range of patterns associated with myelomatosis demonstrates the need for detailed rather than superficial examination of both ultracentrifuge and electrophoresis patterns. Common features may then become apparent.

CALF SERA

During the last few years Dr A. E. Pierce, of the Institute of Animal Physiology, Babraham, Cambridge, has also been collaborating with me in the physico-chemical examination of calf sera with the object of studying in detail and explaining the changes which occur in calf sera with time, and particularly after the ingestion of colostrum. Fig. 8(*a*) contains typical sedimentation patterns for a calf before receiving colostrum. The almost complete absence of a component of sedimentation constant 7 S is noteworthy, a slower component of 4 S and a trace of macroglobulin component only occurring. If the calf is deprived of colostrum, a very slow and variable formation of 7 S material occurs with time (Fig. 8(*b*)). On the other hand, within 16 hours after ingestion of colostrum, the serum pattern changes markedly (Fig. 8(*c*)), the 7 S component now reaching a level more than comparable with that of the adult animal. Fig. 8(*d*) contains sedimentation patterns of colostrum after clarifying by high-speed centrifuging. The occurrence of a component of sedimentation constant 7 S is clear and its superficial resemblance to the 7 S component of post-colostral sera is confirmed by detailed study. It cannot be said that these components are identical, but undoubtedly the immune lactoglobulin (for it proves to carry the antibody activity of the colostrum) enters and leaves the gut and eventually reaches the blood vessels with minor, if any, changes. On the other hand, the much slower sedimenting lactoglobulin and lactalbumin (3 S and slower) do not reach the serum in appreciable quantities. Whether the latter components are held back by the gut which must therefore be specifically permeable to the immune lactoglobulin, or whether, as seems more likely, they penetrate the gut but are otherwise disposed of to prevent appreciable appearance in the serum, are questions which are currently being considered.

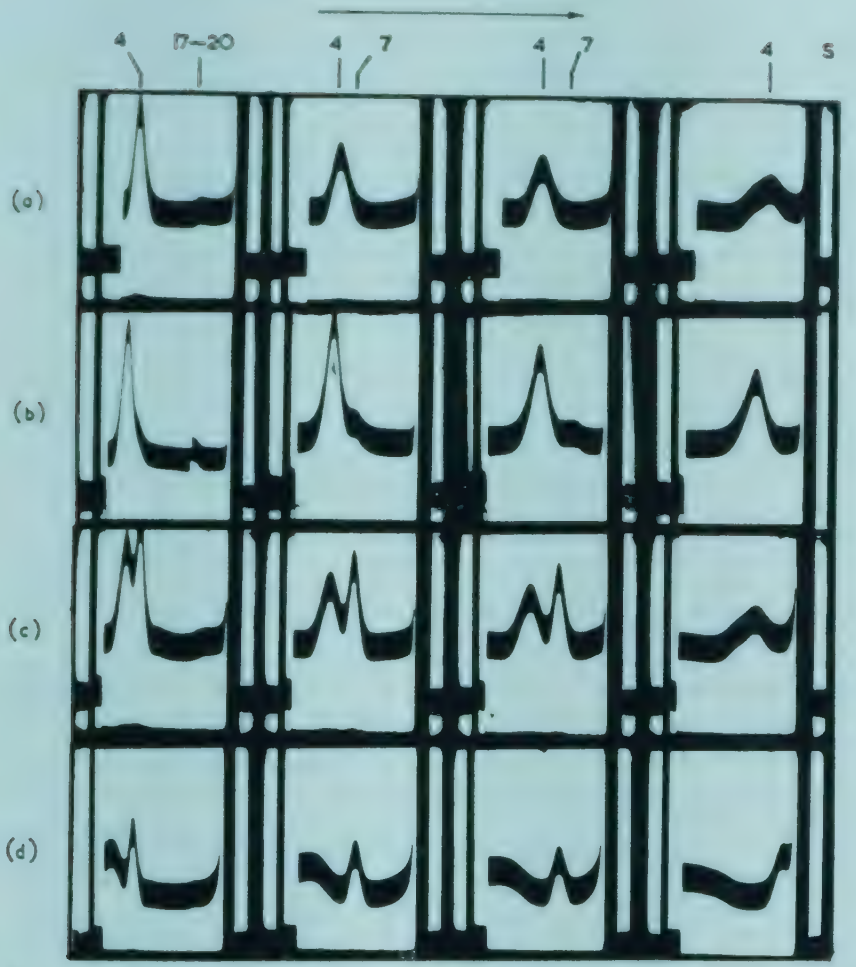


Figure 8. Sedimentation patterns at 59,780 r.p.m. in phosphate buffer (I 0·2, pH 8) for:
a. Serum from newly-born calf which had been deprived of colostrum.
b. Serum from 22-day-old calf which had been deprived of colostrum.
c. Serum from 1-day-old calf which received colostrum 16 hours earlier.
d. Clarified colostrum.

MUSCLE PROTEINS

The mechanism of muscular contraction has long stimulated thought and much work is now being devoted to its many aspects. For many years, in my laboratory, muscle proteins have been studied and recently the subject has been re-opened by Mr A. J. Rowe. It is usually accepted that myosin and actin are the two chief protein constituents in skeletal muscle and that their interaction gives the complex actomyosin, perhaps the chief structural element of muscle. Our studies have therefore been concerned firstly with the single components, since it was thought necessary to understand thoroughly their main features before embarking

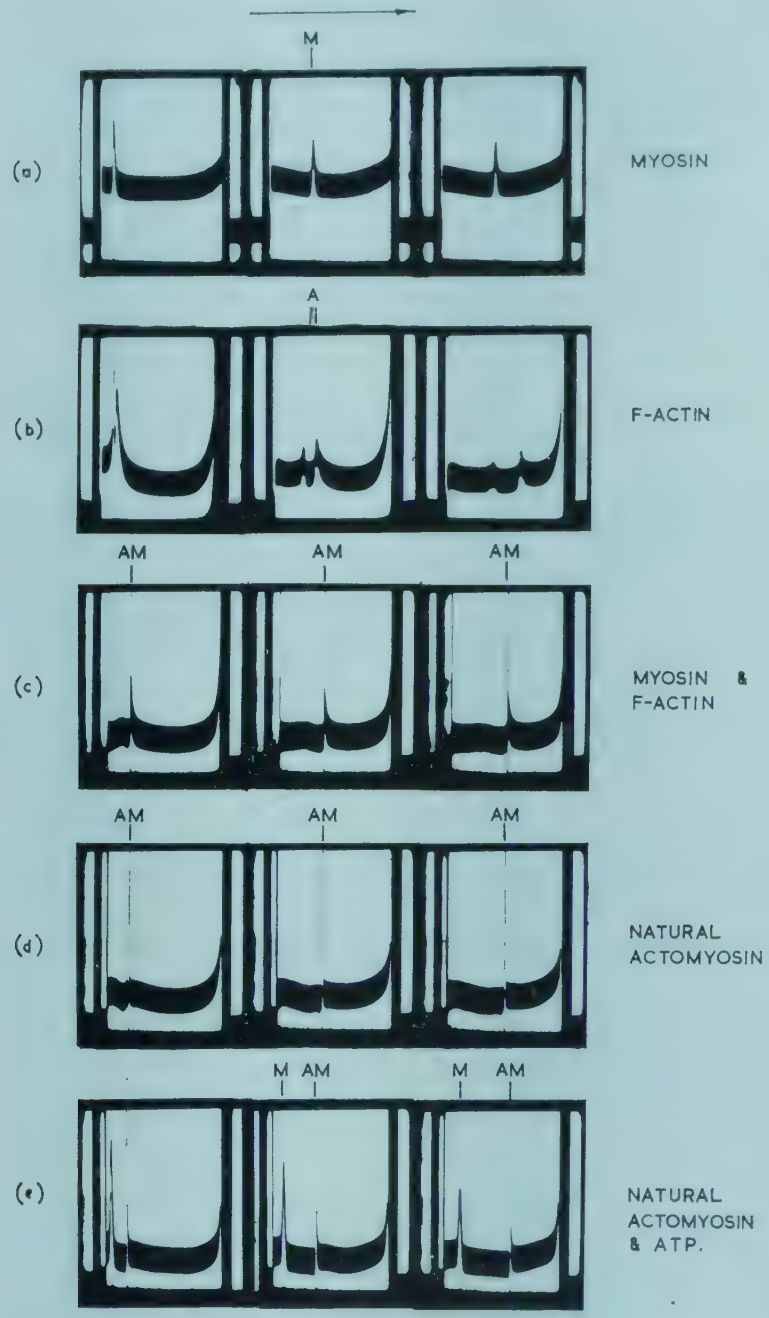


Figure 9. Sedimentation patterns at 16–20° C. in KCl/phosphate buffer (I = 0·6, pH 6·7) at (*a*) 59,780 r.p.m. and (*b–e*) 44,770 r.p.m. for rabbit muscle proteins.

a. Myosin, 0·15 gm./100 ml.

b. F Actin, 1·1 gm./100 ml.

c. Myosin (0·54 gm./100 ml.) + Actin (0·31 gm./100 ml.).

d. Natural actomyosin, 0·52 gm./100 ml.

e. Natural actomyosin, (0·45 gm./100 ml.) + $2 \cdot 10^{-4}$ M adenosine triphosphate.

upon more complicated systems. It should be mentioned that this system is much more difficult than those already mentioned, complicating features occurring consistently even in the study of the individual components.

Fig. 9(*a*) shows the sedimentation of fresh rabbit myosin (M) as a single well-defined and 'unnaturally' sharp peak on a base-line whose curvature is due to the high salt concentration of the buffer used. A plot of s v. c for myosin proves to possess a considerable slope (Fig. 6(*b*)) so that extrapolation to $c = 0$ is of particular importance for further calculation. Such a slope also introduces one of the difficulties of the method. Since sedimentation is slower at the high than at the low concentrations, it can be seen from Fig. 4(*a*, ii) that the advancing edge of the boundary will move more slowly than the trailing edge. In other words, the boundary becomes sharpened as it moves, the extent of the effect increasing with increasing total concentration; it can be so pronounced that the spreading expected on the grounds of polydispersity as well as of diffusion can be completely prevented.

Fig. 9(*b*) shows on the other hand, the sedimentation of actin (A) with at least two different and more rapidly sedimenting species. On mixing myosin and actin, a pronounced viscosity increase was shown to occur by Szent-Gyorgyi and, associated with this increase, it is possible by quantitative measurements to demonstrate the occurrence of new, highly concentration-dependent, sedimenting species AM (Fig. 9(*c*)). Natural actomyosin complex can be extracted as such from the muscle (Fig. 9(*d*)) and it closely resembles the synthetic product particularly in its susceptibility to the action of adenosine triphosphate (ATP). Fig. 9(*e*) shows the partial breakdown of the complex in the presence of ATP, detectable readily in the increased occurrence of the myosin-like sharp peak sedimenting behind. The main features of our results are in good agreement with Szent-Gyorgyi's findings but much remains to be done before we can have a thorough understanding of the interaction between myosin and actin. However, such a goal is well worthwhile, since it may make possible an understanding of the behaviour of a major structural element in the muscle itself.

Recent Developments and Future Possibilities

SYNTHETIC BOUNDARY CELL

With very low molecular weight substances, the separation of the boundary from the meniscus region is very slow. In the synthetic boundary

cell, solution only partially fills the cell and it is so devised that, at moderate speed, a small volume of solvent flows in and forms a layer on the solution. The motion of the boundary thus formed away from the meniscus can be observed immediately. The sedimentation of sucrose has thus been followed.

PARTITION CELLS

Two main types of cell are now available by which the separation of small quantities of differently sedimenting components, suitable for physico-chemical and biological testing, can be achieved.

Various optical and mechanical devices are being tested, and introduced by which, for example, alternative methods of computation become possible, and several experiments can be conducted simultaneously. Undoubtedly the ultracentrifuge of the future will be of the 'press-button' type, the operator feeding in his sample and returning later to find a record of the number of differently sedimenting components, their concentrations and possibly even a description of their specific properties.

REFERENCES

Alexander & Johnson (1949) *Colloid Science*. Oxford.
Edsall (1953) *in The Proteins* edited by Neurath & Bailey. Academic Press, Vol. I, Part B, p. 549.
Schachman (1959) *Ultracentrifugation in Biochemistry*. Academic Press.
Svedberg & Pedersen (1940) *The Ultracentrifuge*. Oxford.

X-RAY STUDIES IN BIOLOGICAL RESEARCH

D. C. Phillips

1. General Outline of the Method

The technique which I have to describe has been covered in a large number of books such as the three volumes of *The Crystalline State* by Bragg, James and Lipson & Cochran and those by Bunn and Robertson which are more particularly written for non-physicists. In this short introduction I shall follow W. L. Bragg in emphasizing the close correspondence between X-rays and visible light and the ways in which they are used to examine the structure of matter in fine detail. In fact the X-ray diffraction technique can be described quite simply as an extension of ordinary microscopy. By its use we are enabled in a sense to see atoms and the ways in which they are arranged in space. Let us then begin by considering what happens when a simple diffraction grating is examined in an ordinary laboratory microscope.

(*a*) Diffraction theory of microscopic vision

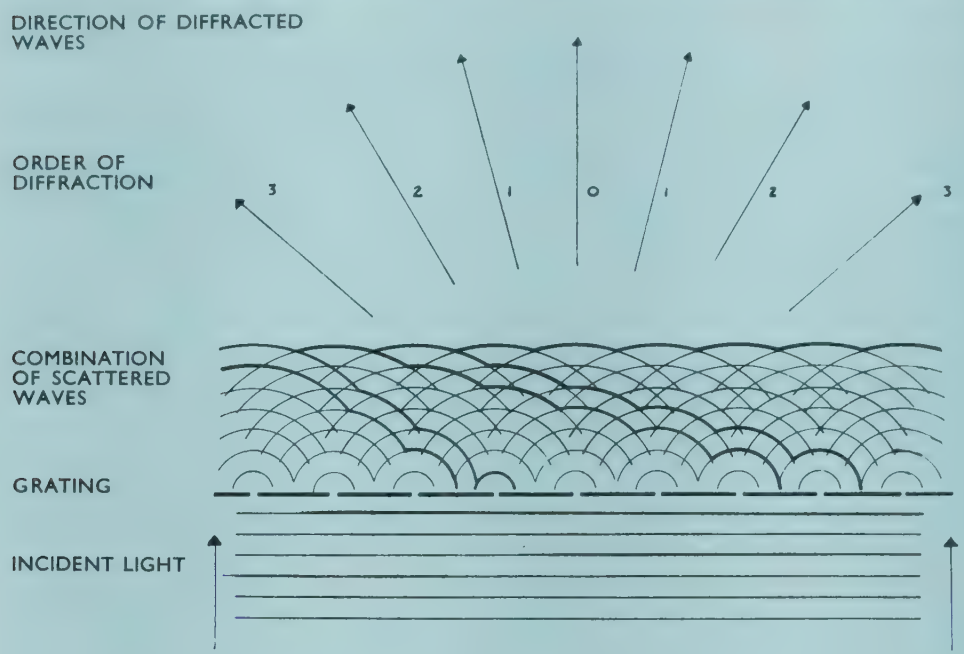
The diffraction grating consists of a large number of parallel lines ruled close together on the surface of a glass plate. For simplicity we can suppose that the ruled lines are all the same width and are perfectly opaque while the spaces between them, which are also constant in width, are transparent. Let us suppose that such a grating is illuminated from behind by a parallel beam of monochromatic light consisting of a train of plane light-waves, and that it is examined from the front by means of a microscope.

The formation of the microscope image takes place in two stages. In the first stage the light-waves are scattered by the grating in a peculiar way. Each transparent line acts as a source of light-waves which radiate from it in the way illustrated in Fig. 1(*a*) and these scattered waves combine with one another to produce new trains of plane-waves travelling in a number of definite directions. It is this co-operative combination of scattered waves which is known as diffraction and it is very important to notice at the outset that each line in the grating contributes to each of the diffraction spectra, as the various trains of diffracted waves are called.

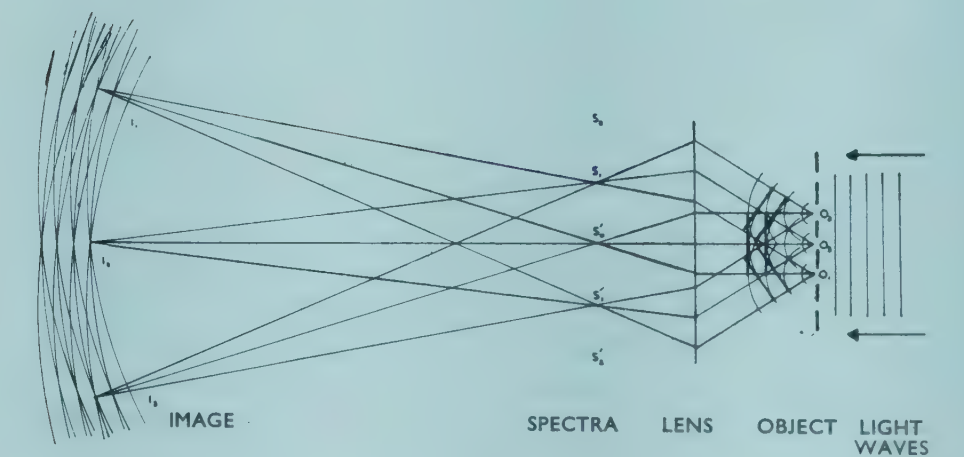
The directions in which the diffraction spectra appear are determined by the wave-length λ of the incident light and by the spacing d, of the grating lines according to the rule

$$\sin \phi_n = n\lambda/d,$$

where ϕ_n is the angle which the n^{th} order diffracted wave makes with the



(1a)



(1b)

Figure 1. Diagrams illustrating (*a*) the diffraction of light-waves by a diffraction grating and (*b*) recombination of the diffracted waves to form an image in an optical microscope.

incident waves and n is a whole number. Clearly the deviation of the diffracted waves from the direction of the incident light increases if the spacing of the lines is decreased but decreases if the wave-length of the light is decreased.

In the second stage of image formation the lens collects the diffracted waves and focuses them in its focal-plane at S, S, S . . . as shown in Fig. 1(*b*). The light-waves, however, pass on through the focal-plane and recombine to form an image of the grating, I_1, I_2, I_3 . . . in the image plane of the objective.

These phenomena can be observed quite easily in a laboratory microscope. The grating is set on the microscope stage and illuminated by

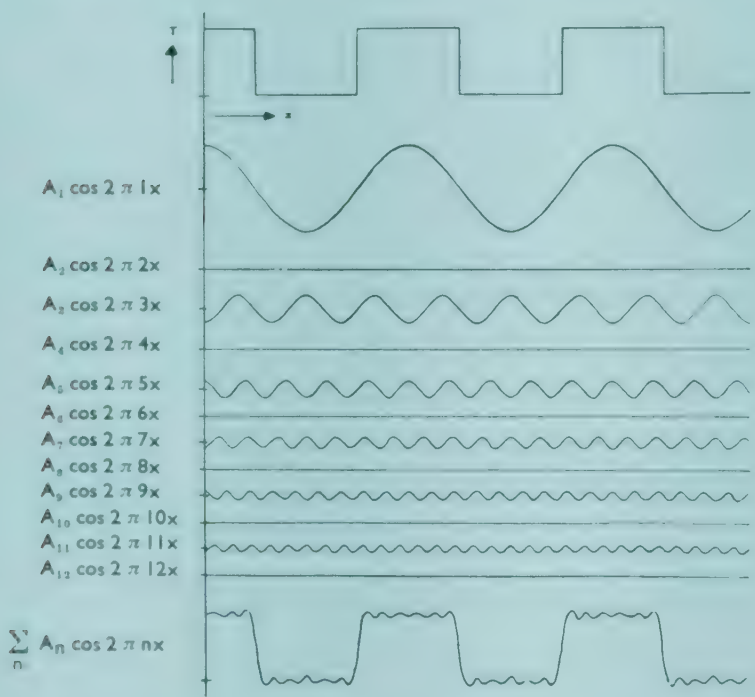


Figure 2. Analysis of the variations of transparency in a grating into harmonic components and the recombination of a limited number of these harmonic components to form an imperfect image of the grating. Notice that some of the harmonic components have zero amplitude because of the symmetry of the grating in which the transparent and opaque sections are equal.

parallel, monochromatic light. The microscope then can be focused in the usual way to give an image of the grating and, if the eyepiece of the microscope is removed, the focused diffraction spectra can be seen in the focal-plane of the objective. The problem now is to determine the relationship between the diffraction spectra and the microscope image.

This relationship is demonstrated most clearly by means of some simple experiments first described by A. B. Porter. Suppose that we place a stop in the focal-plane of the objective so that any or all of the diffracted waves which are focused there can be prevented from contributing to formation of the image. It is found that when some of the diffracted waves are stopped the image is very much changed. Thus if only the central wave or zero-order diffracted wave focused at *S* is allowed to pass, no image of the grating is formed at all and only uniform illumination is observed. If only the zero and first-order diffracted waves are allowed to pass, a hazy image is formed in which the lines have the same spacing as those in a true image of the grating but are blurred out in a simple sinusoidal variation of light intensity. The zero and second-order waves on the other hand also give rise to a simple sinusoidal variation in light intensity but now the image has twice as many lines as the true image. The third-order spectrum gives rise to an image with three times the correct number of lines and the higher-order spectra give images with the apparent number of lines always increasing with the order of the spectrum. In short it appears that the diffraction spectra constitute a harmonic analysis of the variations in transparency in the grating which is analogous to the analysis of a musical note into its fundamental frequency and harmonics. Clearly the image observed in the microscope depends critically on the number of harmonic components which are combined in its formation.

This process of image formation is illustrated in Fig. 2 which shows some of the harmonic components into which the variations in transparency in a simple grating can be analysed together with the result obtained by combining them to form an image.

(*b*) THE USE OF X-RAYS IN MICROSCOPY

We can now consider the problems involved in looking at the structure of matter in atomic detail. Consider again the effect of reducing the distance between adjacent lines in the diffraction grating. The first-order spectrum, and of course all the higher orders too, will deviate more and more from the direction of the incident light until they can no longer be collected by the objective lens. Only the central wave will then reach the image plane and, as we have seen, no image will be formed. This situation will be reached with certainty in the arrangement shown in Fig. 1(*a*) when the spacing of the lines is the same as the wave-length of the incident light. Clearly the way to regain resolution is to decrease the

wave-length of the incident radiation so as to bring the diffracted waves again within the aperture of the objective. This approach has, of course, been used in the design of ultra-violet microscopes but even such instruments cannot resolve the individual atoms which are only a few Angström units (10^{-8} cm.) apart, about one-thousandth part of the wave-length of light. In order to resolve atoms we must use a radiation with a comparable wave-length: X-rays (or electrons or neutrons) must be used.

Here we encounter a fundamental difficulty: X-rays cannot readily be focused to form an image. We are forced then to face a conscious separation of the two steps in image formation which we have been considering: we have to measure the X-ray diffraction spectra of the object under investigation and then build up the image of it from a knowledge of the amplitudes of its harmonic components.

Stated in this way the process sounds quite straightforward but there is unhappily one major difficulty. We can determine the amplitudes of the harmonic components needed to make up the image but not their relative phases. The importance of the phases is evident in Fig. 2 where it can be seen that some of the component waves have peaks at the centres of a maximum in the image while others have troughs. In mathematical symbols we see that the variation in transparency across the grating is given as a function of position x by the well-known Fourier series,

$$T(x) = \sum_n A_n \cos 2\pi nx \qquad (1)$$

where the coefficients A_n have to be known in amplitude (which can be observed) and in sign (which cannot be observed) if a true image is to be obtained.

In practice the situation is often more difficult than that shown in Fig. 2. The grating illustrated there has a centre of symmetry at the middle of each line so that the component waves must also have centres of symmetry at the middle of each line, that is, each wave must have either a peak or a trough at this same point in the image. But if the object is not symmetrical in this way—and most biologically interesting structures are not—the component waves are not restricted in phase and may be in or out of step with one another by any amount. Image formation then is governed by the Fourier series:

$$T(x) = \sum_n A_n \cos (2\pi nx - \alpha_n) \qquad (2)$$

in which the amplitudes A_n can be measured but the phase angles, α_n,

cannot be measured and have to be determined by some indirect method. This is the famous 'phase-problem' of X-ray crystallography.

(*c*) DIFFRACTION BY CRYSTALS

This use of the common title for the technique brings us now to a consideration of the objects which are most profitably studied by X-ray diffraction. Crystals are formed by the regular arrangement of atoms or molecules and can be thought of as a kind of three-dimensional wallpaper in which the unit pattern, the unit-cell, is regularly repeated in space. The

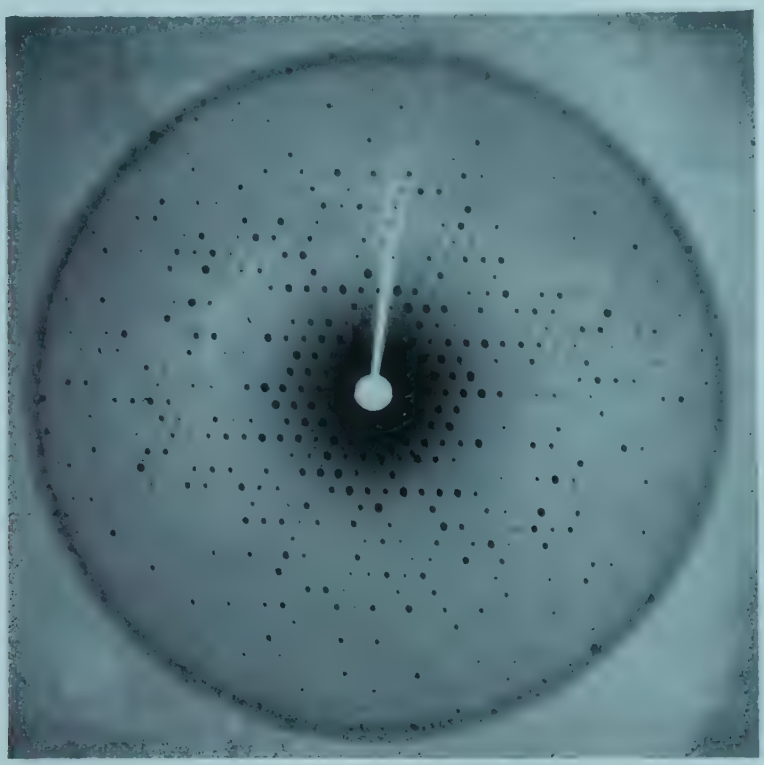


Figure 3. X-ray diffraction photograph from a single crystal of haemoglobin. The different intensities of the spots provide the information from which the structure is utimately deduced (cf. Fig. 6: Perutz, M. F. (1958) *Endeavour*, XVII, p. 190).

unit-cell may contain one or more symmetrically related molecules but the possible symmetries of the arrangement, which are discussed in all the standard accounts of the subject, need not concern us. It is enough to appreciate that a crystal may be regarded as a three-dimensional diffraction grating capable of diffracting X-rays.

The diffraction spectra due to two-dimensional or crossed gratings are

quite familiar. Examples can be seen by looking at a distant light through handkerchief or umbrella fabric when a two-dimensional array or pattern of spectra appears. Similarly the complete X-ray diffraction pattern due to a three-dimensional grating, a crystal, is three-dimensional also. This diffraction pattern cannot be observed completely in one experiment but sections of it can be observed such as the section of a typical pattern shown in Fig. 3. Such sections clearly are very similar to the diffraction patterns of simple crossed-gratings: they correspond to two-dimensional views or projections of the three-dimensional crystal structure. The complete diffraction pattern, or array of spectra, is built up from the patterns due to all possible views of this kind. The complete pattern of which Fig. 3 is a part, for example, comprises many parallel sheets, each with the same geometrical arrangement of points, equally spaced along a line normal to the section shown.

The X-rays are scattered by the electrons in a crystal. Following the optical analogy outlined above, we see therefore that the amplitude of each X-ray diffraction spectrum tells us the amplitude of a harmonic component or wave in the distribution of electrons in the crystal. The position of the spectrum in the diffraction pattern tells us the direction and wave-length of this wave. Measurement of the diffraction spectra therefore gives us all the information that is needed to build up an image of the structure—*except the relative phases of the waves.*

(*d*) CALCULATIONS

The distribution of electrons in a crystal is given by a Fourier series of the same kind as equation (2) modified only to give the distribution in three-dimensions. Thus:

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l F(hkl) \cos\{2\pi(hx + ky + lz) - \alpha(hkl)\} \qquad (3)$$

where $\rho(xyz)$ represents the electron density in the crystal at the point x, y, z; V is the volume of the unit-cell of the crystal; the triple summation is over all the diffraction spectra identified by the three-fold indices h, k and l; the quantities $F(hkl)$, known as the structure amplitudes, are directly related to the amplitudes of the diffraction spectra (hkl) and the $\alpha(hkl)$ are the unmeasurable phases.

These phases $\alpha(hkl)$ are the only unmeasurable quantities in this equation so that when they have been determined by some method in any particular investigation equation (3) can be used to synthesize an image of the electron distribution. The electron density is calculated at a number of points in

the unit-cell and the image is usually presented in the form of a contour map in which lines are drawn through all points in the unit-cell with the same electron density. If enough spectra were included in the calculation, individual atoms would appear in the image as concentrations of electrons, that is, as hills in the contour map. Such a map is shown in Fig. 4 which shows a projection in two-dimensions of the structure of codeine hydrobromide dihydrate. The results of calculating the electron density through-

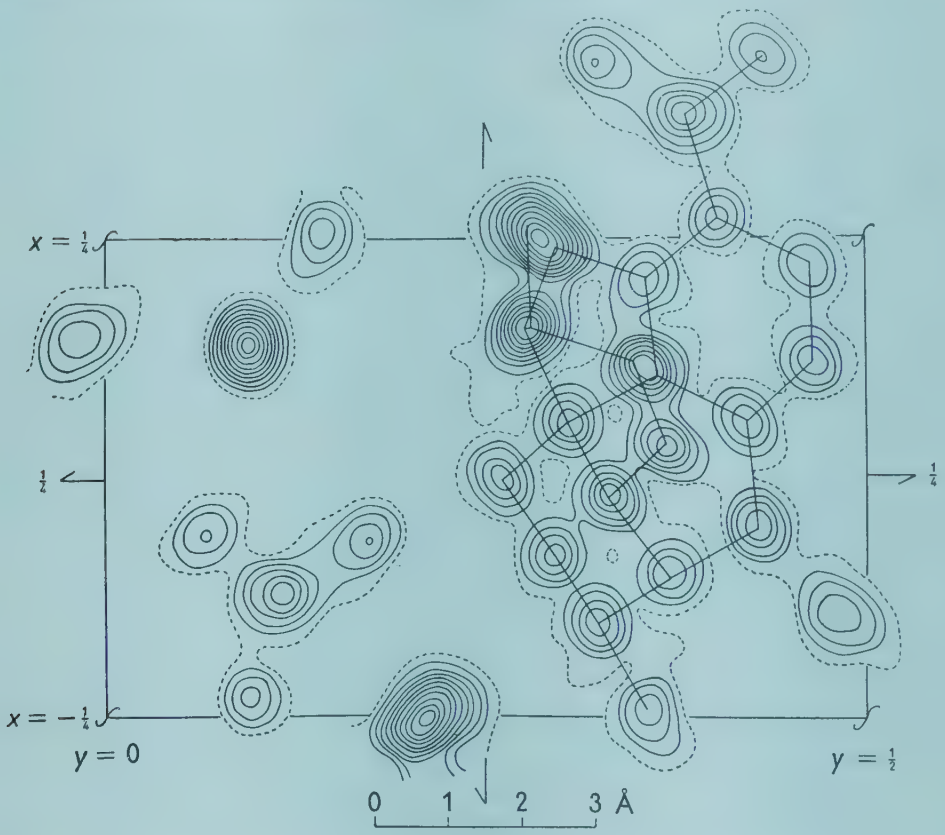


Figure 4. Electron density in codeine hydrobromide dehydrate projected on the (001) face of the unit-cell; contours at intervals of 2e.Å^{-2} (8e.Å^{-2} for the bromine atom), with that at 2e.$^{-2}$Å broken (cf. Fig. 2: Lindsey, J. M. & Barnes W. H. (1955) *Acta, Crystallographica*, **8**, 227).
(Published by Ejnar Munksgaard Ltd., Copenhagen, Denmark.)

out the unit-cell are usually presented in the same way by means of contour lines in sections at various levels in the cell (cf. Fig. 7).

The process of image formation clearly does not involve the use of very advanced mathematics but it is very often difficult to handle and time-consuming because of the large number of diffraction spectra and the large number of points at which the electron density has to be calculated.

Thus, for example, in the latest work on the protein myoglobin about 50,000 diffraction spectra have been measured and the electron density has been calculated at more than 100,000 points in the unit-cell. Such calculations can be undertaken only by the use of the most modern computers and indeed the recent developments of the technique have been very closely connected with the concurrent development of these computers.

(*e*) THE PHASE PROBLEM

Finally in this introduction to the theory of the method, we have to consider ways in which the phase problem may be solved. No general solution has yet been found but several methods have been developed which lead to solutions in particular cases. Of these the most widely used are:

(i) TRIAL-AND-ERROR

The trial-and-error method depends upon the fact that when the natures and positions of all the atoms in a structure are known, the intensities and phases of the diffraction spectra can be calculated. The structure amplitudes, $F(hkl)$, are given by the equation

$$F(hkl) = [\{\sum_j f_j(hkl) \cos 2\pi (hx_j + ky_j + lz_j)\}^2 \cdot \{\sum_j f_j(hkl) \sin 2\pi (hx_j + ky_j + lz_j)\}^2]^{\frac{1}{2}} \quad (4)$$

where the summations now are over all the atoms in the unit-cell which have positions $x_j\, y_j\, z_j$ and scattering powers which are given by the functions $f_j(hkl)$, the atomic scattering factors. The phases $\alpha(hkl)$ are given by the formula

$$\tan \alpha(hkl) = \frac{\sum_j f_j(hkl) \sin 2\pi (hx_j + ky_j + lz_j)}{\sum_j f_j(hkl) \cos 2\pi (hx_j + ky_j + lz_j)}. \quad (5)$$

These formulae represent simply the vector addition of contributions from all the atoms to each particular diffraction spectrum *hkl*. The contribution of each atom depends upon its size and shape (which determine its atomic scattering factor, f, and are well known) and its position in the structure.

The trial-and-error method does not depend upon image formation at all. The atomic arrangement is guessed, usually from its symmetry and the limitations this imposes on the packing of atoms and molecules of

known size, and the intensities of the diffraction spectra corresponding to this postulated structure are calculated (4) and compared with the observed values. Many simple structures have been determined quite exactly by this method and, as we shall see, it is also used in the investigation of quite complex structures.

(ii) THE PATTERSON SYNTHESIS

It can be shown that a synthesis of the type (3) above in which the observable coefficients $|F(hkl)|^2$ are used with all the phase angles zero,

$$P(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l |F(hkl)|^2 \cos 2\pi (hx + ky + lz), \qquad (6)$$

gives information about all the interatomic vectors in a structure. A structure comprising n atoms gives a Patterson function containing n^2 peaks, n of which are superimposed at the origin since they indicate the distance of each atom from itself. These syntheses thus are very difficult to interpret in detail for all but the simplest structures but useful information, such as the relative positions of outstandingly heavy atoms, often can be derived from them.

(iii) THE HEAVY-ATOM METHOD AND FOURIER REFINEMENT

A sufficiently heavy atom in a structure may so overwhelm the contributions of the other atoms to the diffraction spectra that the phases are largely determined by it. The position of such an atom usually can be determined from the Patterson synthesis. The phases of this atom's contribution to the diffraction spectra are then calculated and used as a first approximation to the phases of the observed spectra in a first synthesis of the image. This imperfect image is then improved by a process known as Fourier refinement. Additional atomic positions which appear in the first approximate image of the structure are used in the second set of phase-calculations and these phases are used in turn with the observed structure amplitudes in a second synthesis of the image. The process is continued until no new detail and no changes in atomic positions appear in successive image calculations.

(iv) ISOMORPHOUS REPLACEMENT

This is the most powerful method yet developed and its use is responsible for most of the outstanding results which are now being obtained. Suppose that an atom can be added to each unit-cell in a crystal without

otherwise affecting the structure in any way. The amplitudes of the diffraction spectra from the two isomorphous crystals will differ slightly because of this extra atom's contribution: the position of the extra atom can be determined by use of the Patterson synthesis. Then for a centrosymmetric crystal, for which the phase angles must be 0 or π (cf. Fig. 2), we can calculate the sign of this atom's contribution to each spectrum. Hence, by observing the change in amplitude of each spectrum due to the addition of this atom, we can deduce the sign of the spectrum. Thus, for example, when the extra atom makes a positive contribution to a particular spectrum and the amplitude of that spectrum is reduced by addition of the atom, the sign of that spectrum must be negative.

The situation is more complicated for non-centrosymmetric structures: at least three isomorphous crystals with two different atoms changed or added at different positions in the unit-cell are needed to determine general phases.

2. Experimental Methods

The experimental techniques of X-ray crystallography are concerned with measuring the intensities and directions of the diffraction spectra.

The X-rays most commonly used are copper characteristic $K\alpha$ radiation with wave-length 1·541 Å, generated in X-ray tubes which are operated at about 30 kV. and from 20 to 100 mA. Some other radiations, notably Molybdenum $K\alpha$ are also used. The X-ray beams usually are collimated by pin-hole apertures about 0·5 mm. in diameter. The crystals which have linear dimensions up to about 0·5 mm. are supported in the X-ray beam on fine glass rods or inside thin-walled glass tubes mounted on goniometer heads by means of which they can be oriented accurately. In order to record the diffraction spectra satisfactorily it is necessary to orient a crystal to within about 5′ of arc for rotation about three mutually perpendicular axes.

Hitherto the diffraction spectra have been recorded most often photographically. Each section of the three-dimensional pattern, isolated by means of screens, is recorded on a separate photograph by means of a camera which moves the specimen and the film in appropriate ways. The photograph shown in Fig. 3 is typical. The intensities of the spectra are estimated from the blackness of the spots. Nowadays, however, a very large number of measurements are being made for the determination of increasingly complex structures and a great deal of effort is being put into devising means for measuring the diffraction spectra auto-

matically and at high speed. Some detector other than film is used, usually a Geiger, proportional or scintillation counter, together with means for setting the crystal and the detector to record each separate spectrum in turn. At the Royal Institution, for example, Dr U. W. Arndt and I have devised such an apparatus which measures automatically about 50 spectra an hour and prints out the intensity of each in a form suitable for immediate transfer to a digital computer for the subsequent calculations.

3. Examples of X-Ray Structure Determination

The arrangement of atoms in very many chemical compounds has now been determined by this method. Only a few of the results can be mentioned here but many more have been described in review articles such as those by Kendrew & Perutz, and by Crick & Kendrew, and all are being collected in Structure Reports published on behalf of the International Union of Crystallography. Interatomic distances for most types of chemical combination which have been determined by this and other methods are collected in *Interatomic Distances*.

Nearly all of these results are important to biologists who are increasingly interested in the dependence of function on structure. Progress in this field has been described recently by Perutz in a review which shows very clearly the large part being played by X-ray analysis in the development of the subject. The few examples collected here have been chosen to illustrate the types of structure now being investigated by this method.

(*a*) VITAMIN B_{12}

The structures of very many natural products have been determined by X-ray analysis, often before their chemical constitution was known. The structure of benzyl penicillin (Crowfoot *et al.*) was first determined in this way as were the structures of a number of terpenes, steroids and alkaloids. Outstanding among these analyses is that of vitamin B_{12} by Dr Dorothy Hodgkin and her colleagues. Dr Hodgkin has now described much of this analysis in detail and most of the following remarks are taken from one of her accounts of the work.

Vitamin B_{12} was first isolated in a crystalline form in 1948 and by 1951 it was known that the vitamin had an approximate formula $C_{61-64}H_{83-92}O_{13-20}N_{14}PCo$ and contained a number of well-characterized chemical groups which together made up about half its molecular weight. The structure of the remaining half of the molecule was dis-

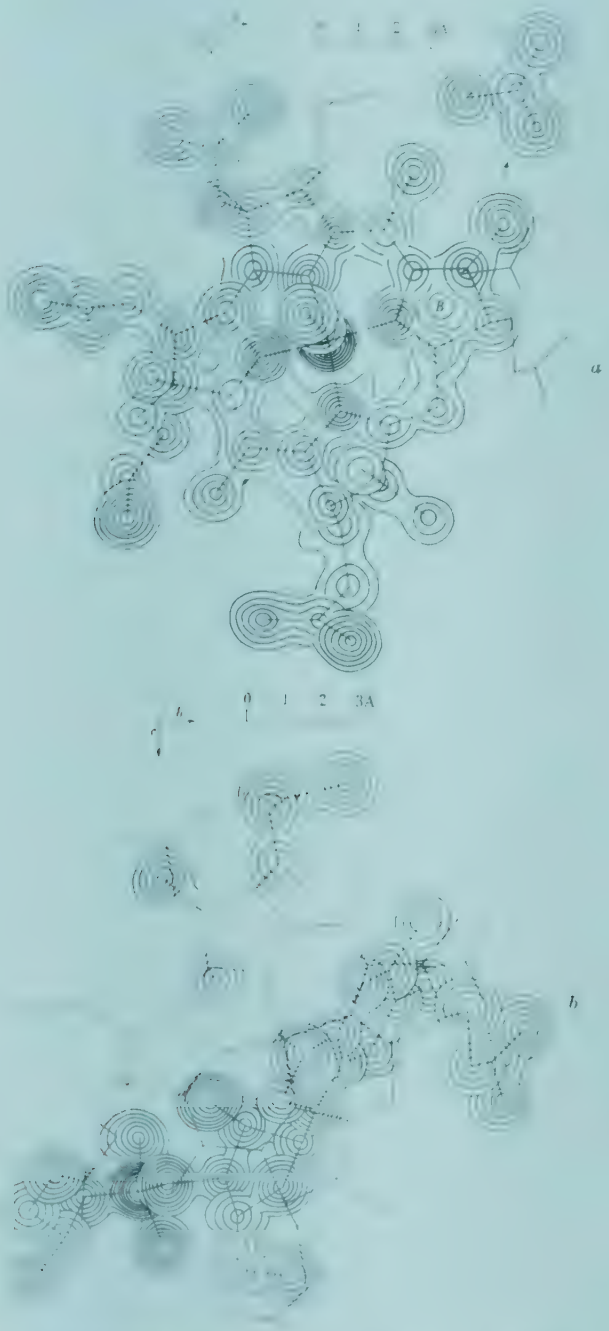


Figure 1. Electron density levels over the atomic peaks in ρ 3 of wet B_{12}. For clarity the structure is shown in two parts: (*a*) The 'planar groups' with the cyanide and most of the side chains; (*b*) the benzimidazole, ribose, phosphate, propanolamine, and the remaining side chains of the planar group (cf. Fig. 16; Hodgkin, D. C. *et al.* (1957) *Proc. Roy. Soc.* A, **242**, 254).

covered very largely by X-ray analysis. This started in 1948 and depended on the study of four different crystal structures, wet and dry B_{12}, a selenocyanide derivative of B_{12} and a hexacarboxylic acid obtained by degradation of the vitamin.

Patterson syntheses (equation 6) were used to find the positions of the heavy cobalt atoms in these structures and the analysis then proceeded, roughly speaking, by parallel Fourier refinements of the various structures starting in each case with phases determined by the heavy atoms alone. This short statement conceals a fascinating story, the most amazing part of which was the way in which different parts of the molecule were recognized one by one among the false details which were present in the early approximate images of the structures.

The result now is that the atoms (other than hydrogen) in this very large molecule have been located to within less than 0·5 Å in two different crystal structures. In addition the absolute configuration of the molecule is known and so is the exact stereochemistry of all the different asymmetric centres present. These results are illustrated in Fig. 5 which shows the X-ray image of wet vitamin B_{12}.

(*b*) PROTEINS

Proteins have been under investigation by X-ray methods now for many years. Two main lines of attack have been used. The first is the indirect method of determining as accurately as possible the structures of the amino-acids from which proteins are known to be built. Many beautiful analyses of these structures have been completed, mostly by Pauling, Corey and their colleagues at the California Institute of Technology. They provide striking examples of the accuracy attainable in X-ray structure determination by means of which interatomic distances in relatively simple organic structures such as these can be measured to within 0·01 Å.

The object of the work on simple peptides was to discover the general arrangement of atoms associated with the peptide link in the hope that structural principles which determine the configuration of the polypeptide chain in proteins would be made evident. It did in fact lead to the suggestion by Pauling, Corey & Branson that the polypeptide chain in many proteins exists in a helical configuration with certain definite dimensions. It is now almost certain that their α-helix, shown in Fig. 6, is a basic element in the structure of many fibrous proteins and that it is present also in the structure of globular proteins.

This indirect approach to the structure of globular proteins has, however, been superseded by the direct determination of their structures. Despite many years' work it was not until 1953 that complete success first appeared possible. In that year Dr M. F. Perutz and his colleagues at the Medical Research Council Unit in Cambridge working on haemoglobin showed that heavy atoms could be attached to protein molecules at

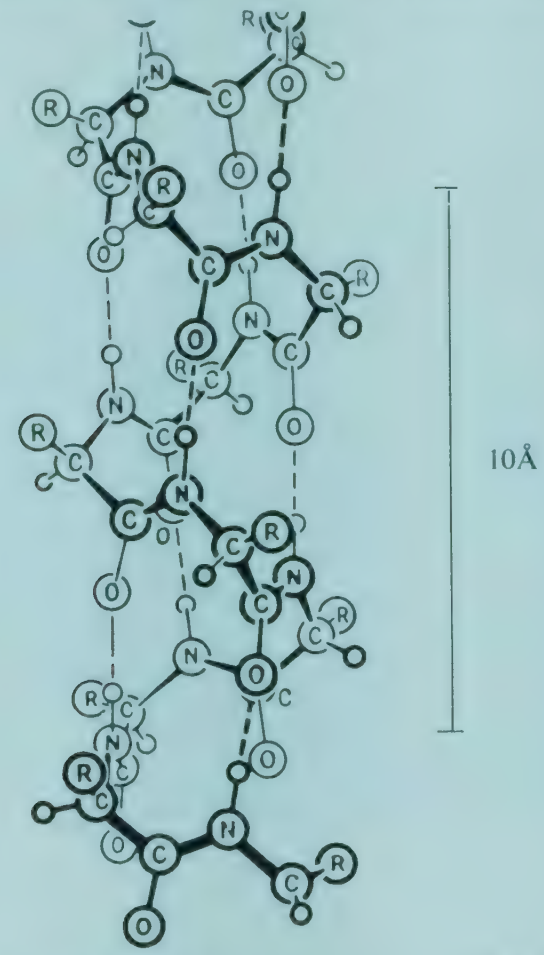


Figure 6. The α-helix of Pauling, Corey & Branson. The atoms marked R represent side-chains (cf. Fig. 17: Perutz, M. F. (1958) *Endeavour*, XVII, 190).
(Permission to reproduce from Pauling, Corey & Bransom, reference 17.)

specific sites in the crystal structure and that the resultant complexes gave diffraction patterns which were sufficiently different from that of the native protein for the method of isomorphous replacement to be used to determine the phases of the diffraction spectra.

The method was used immediately (Bragg & Perutz) to determine phases for the synthesis of a projection of the structure. Unfortunately

however there was so much overlapping of atoms in this projection that very little information could be derived from it and it was clearly necessary to extend the work to three dimensions. This extension of the work on haemoglobin therefore was begun and at the same time the method was applied to the study of myoglobin, a closely related protein, by Dr J. C. Kendrew and his colleagues also at the M.R.C. Unit in Cambridge and at the Royal Institution in London. The function of haemoglobin is to carry oxygen from the lungs to all parts of the body, where myoglobin

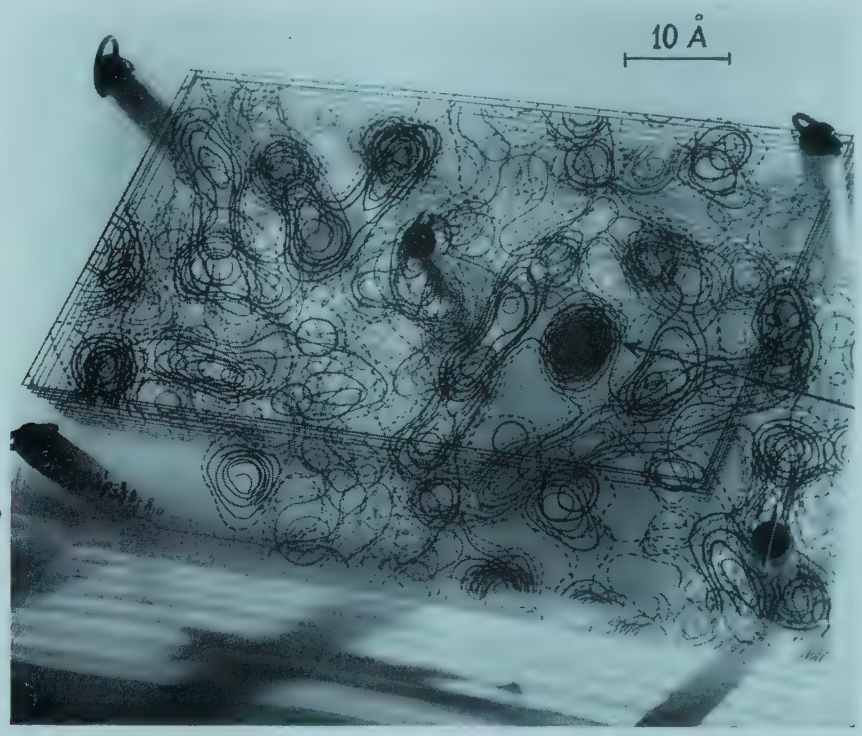


Figure 7. Three-dimensional map showing the distribution of electron density along a series of parallel sections through two neighbouring molecules in a crystal of myoglobin. The arrow points at high peak representing the iron-containing haem group (after J. C. Kendrew) (cf. Fig. 11: Perutz, M. F. (1958) *Endeavour*, XVII, 190).

holds some of it in store. Myoglobin, with molecular weight 17,000 comprising about 1,200 atoms which might be 'seen' by X-rays, is the smaller molecule; it is one-quarter the size of haemoglobin and the analysis of its structure has proceeded so far in three stages. In the first stage (Bluhm *et al.*) a two-dimensional projection of the structure was determined but again, as with haemoglobin, there was so much overlapping of atoms in the projection that no information about the structure

of the molecule could be derived from it. The second stage (Kendrew *et al.*) was to obtain an image of the structure in three dimensions but with limited resolution. Only the 400 lowest-angle diffraction spectra, corresponding to a resolution of 6 Å, were included in this stage of the investigation but it must be remembered that these spectra had to be measured from crystals of the native protein and from at least two isomorphous crystals containing different heavy atoms in different positions. In fact measurements from six isomorphous crystals were used to establish the phases.

The resolution aimed at in this stage of the analysis was that calculated to show α-helices, or similar structures, if they exist in the molecule, as rods of high electron density. A view of the image actually obtained is shown in Fig. 7. In addition to one region of outstandingly high density which is taken to represent the iron-containing haem group, the part of the molecule to which oxygen attaches itself, the image does indeed contain a number of prominent rods of high electron density. These generally run fairly straight for up to 40 Å but there are some curved ones and the straight segments often are joined by fairly sharp corners. It seems certain that these rods represent the polypeptide chain itself and their appearance suggested that the straight segments at least might be in the α-helix configuration. The various rods in the structure are however intertwined in a very complex manner as shown in Fig. 8, which is a photograph of a model of the regions of highest electron density.

The third stage of the analysis which is now nearing completion has been directed towards increasing the resolution so that more details of the molecular configuration can be recognized. It seems quite likely that eventually individual atoms may be seen in this structure but all the available diffraction spectra, about 20,000 in all, then will have to be included in image formation. This presents considerable technical difficulties so that the third stage has been limited to including 10,000 spectra to give 2 Å resolution. Again these spectra have had to be measured for the native protein and a number of isomorphous crystals containing heavy atoms. This time five crystal types in all were used so that 50,000 spectra were measured.

This third stage of the work on myoglobin and the first three-dimensional work on haemoglobin, aimed at producing an image with 6 Å resolution, have both come to fruition in the past few weeks. The images are now being analysed and fresh details are recognized almost daily so that a complete account cannot yet be given. However, it can be said that parts of the polypeptide chain in myoglobin appear to be in the α-helix

configuration. The image of haemoglobin on the other hand looks very like that of myoglobin at the same resolution (Fig. 7). We can be sure that the architecture of these very complicated molecules will soon be known in atomic detail.

Figure 8. Model of the myoglobin molecule. The white sausage represents the polypeptide chain; the grey disc is the haem group. The little balls represent the positions of different heavy atoms used for determining the structure. The scale is in Angström units (cf. Fig. 13(*a*): Perutz, M. F. (1958) loc. cit.).

(*c*) FIBRES

Many structures which are important biologically cannot be studied in single crystals so that we must consider, if only briefly, to what extent X-rays can be used in the study of imperfectly crystalline material.

The theory of the microscope with which we began can be adapted quite easily to cover the examination of objects which are not, like diffraction gratings or crystals, built up by the regular repetition of a unit

pattern. The difference lies in the nature of the diffraction pattern. The diffraction pattern produced in X-rays by a single molecule for example is a continuous distribution of X-ray intensity and in fact if this pattern could be observed, the molecular structure could be derived from it more readily than it can be derived from the diffraction pattern of a crystal. Unfortunately, however, the X-ray intensity diffracted by such a specimen would be too weak to measure even if the specimen could be handled and we are lucky that crystals so often exist in which many molecules are accurately aligned and which are capable of diffracting measurable intensities. The effect of arranging the molecules regularly in a crystal structure is to concentrate the scattered X-ray intensity in particular directions which are determined by the dimensions of the crystal lattice. The continuous diffraction pattern of a single unit-cell is observed only in certain discrete directions in the diffraction pattern of a crystal.

Specimens which are imperfectly crystalline give diffraction patterns in which the scattered X-rays are not confined to discrete directions in this way but spread more generally in all directions. Thus the pattern due to a random arrangement of molecules can be thought of as arising from the combination of diffraction patterns from molecules in a large number of different orientations. The only detail left in such a pattern is a radial variation in intensity and this provides information only about average interatomic distances in and between the molecules. X-ray diffraction has in fact been used in this way in the study of liquids and gases, and even in an attempt to identify the main features of protein structure (Arndt & Riley), but it is clearly not a very powerful technique. It is obvious that X-ray diffraction is most useful in the study of structures or parts of structures which repeat regularly in space. Fortunately regularities of this kind abound in biological material.

The most important imperfectly crystalline specimens which have been studied by X-ray diffraction are fibres. The best of these, from an X-ray point of view, can be thought of as collections of small crystallites generally embedded in a certain amount of amorphous material. In a well-oriented fibre the crystallites all lie with one axis, the fibre axis, parallel to the length of the fibre but the orientations around this axis are random. The diffraction pattern of such a specimen then is similar to that which would be obtained from a single crystal if it were rotated continuously about one axis. Even for the best oriented fibres therefore it is generally impossible to isolate the individual diffraction spectra arising from separate crystallites or to use them in synthesizing an image of the structure.

Fibre diagrams, as photographs of these diffraction patterns are called, in practice are often less detailed even than the foregoing description would suggest. A fibre often contains no crystallites in which the molecules are regularly arranged in three dimensions but is built up rather from long polymer molecules which all lie approximately parallel to the fibre axis but are not regularly related to one another in any other way. The fibre diagram of such a specimen shows no discrete spectral spots but only continuous variations of intensity broken up into layers by the regularity in the direction of the fibre axis. Such a diagram is shown in Fig. 9 (Cowan *et al.*).

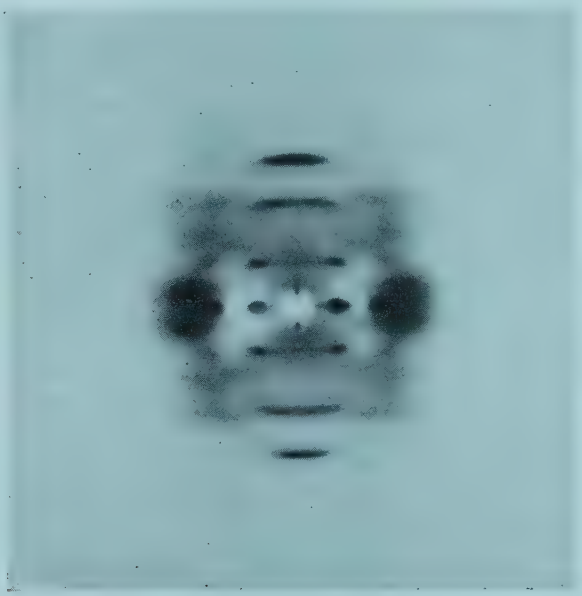


Figure 9. X-ray diffraction pattern of rat tail tendon stretched by 8 per cent (cf. Cowan, P. M., North, A.C.T. & Randall, N. T. (1955) in *Fibrous proteins and their Biological Significance*. Cambridge Univ. Press.) reproduced by permission of The Company of Biologists Ltd.

Nevertheless fibre diagrams have given invaluable information about the structures of many compounds, particularly the fibrous proteins, synthetic polypeptides and deoxyribonucleic acids (DNA).

These recent successes have depended very much on the realization that helical arrangements of atoms are the basis of fibre structures. The diffraction theory of these arrangements was worked out in 1952 and since then it has been found that the helical nature of a fibre structure often can be recognized from its diffraction pattern at a glance—by an expert who knows what he is looking for.

Although the helical nature of the atomic arrangement in a fibre and its general parameters often can be established with some certainty from its fibre diagram, the exact positions of the atoms are not so readily found. In fact, since an image cannot be synthesized, the positions of the atoms can be found only by the process of trial-and-error. Scale models are built to discover atomic arrangements which are consistent with the known stereochemistry of the component parts of the structure and the diffraction patterns corresponding to these models are calculated and compared with the observed fibre diagrams. Clearly the results obtained in this way can never be as convincing as those obtained by the study of single crystals, but nevertheless several important structures derived by a very careful application of this method are now generally accepted. Thus for example it appears certain that the configuration of the α-polypeptides (Bamford *et al.*) is based on the α-helix which is also found in the α-keratin-type fibrous proteins. The work on DNA must also be mentioned. The model of this structure proposed by Watson & Crick was inspired in part by the fibre diagrams obtained by Wilkins, Seeds & Wilson and by Franklin & Gosling. Further analysis of these fibre diagrams together with extensive model building, particularly by Wilkins and his colleagues at the M.R.C. Unit, King's College, London, has shown that the diffraction patterns are fully consistent with the double-helix structure proposed by Watson & Crick, though some features of the initial model needed modification. Wilkins and his colleagues have also shown with this X-ray method that the same structure is present in live sperm heads (Feughelman *et al.*).

(*d*) VIRUSES

Valuable as the results from the study of fibres have been, it is clear that the full power of the X-ray method can be developed only when good single crystals are available. Furthermore as we have seen the most successful work is done when the crystalline material can be modified, for example by the addition of heavy atoms, to suit the requirements of the technique. This is perhaps a good place to appeal for co-operation in this part of the work. Anyone who wants to know the arrangement of atoms in biologically important material should be on the look-out for crystals. They are in fact more common than one might suppose, as is perhaps shown most clearly by the existence of virus crystals.

The virus studied most extensively by X-rays is the rod-shaped, tobacco mosaic virus, TMV. True crystals of TMV with sharp faces are

found within the cells of the host plant but they are too small to be examined with X-rays. Unfortunately no one has yet succeeded in producing true crystals from extracted virus particles but these particles easily form paracrystalline gels, in which the virus rods are all parallel and in hexagonal

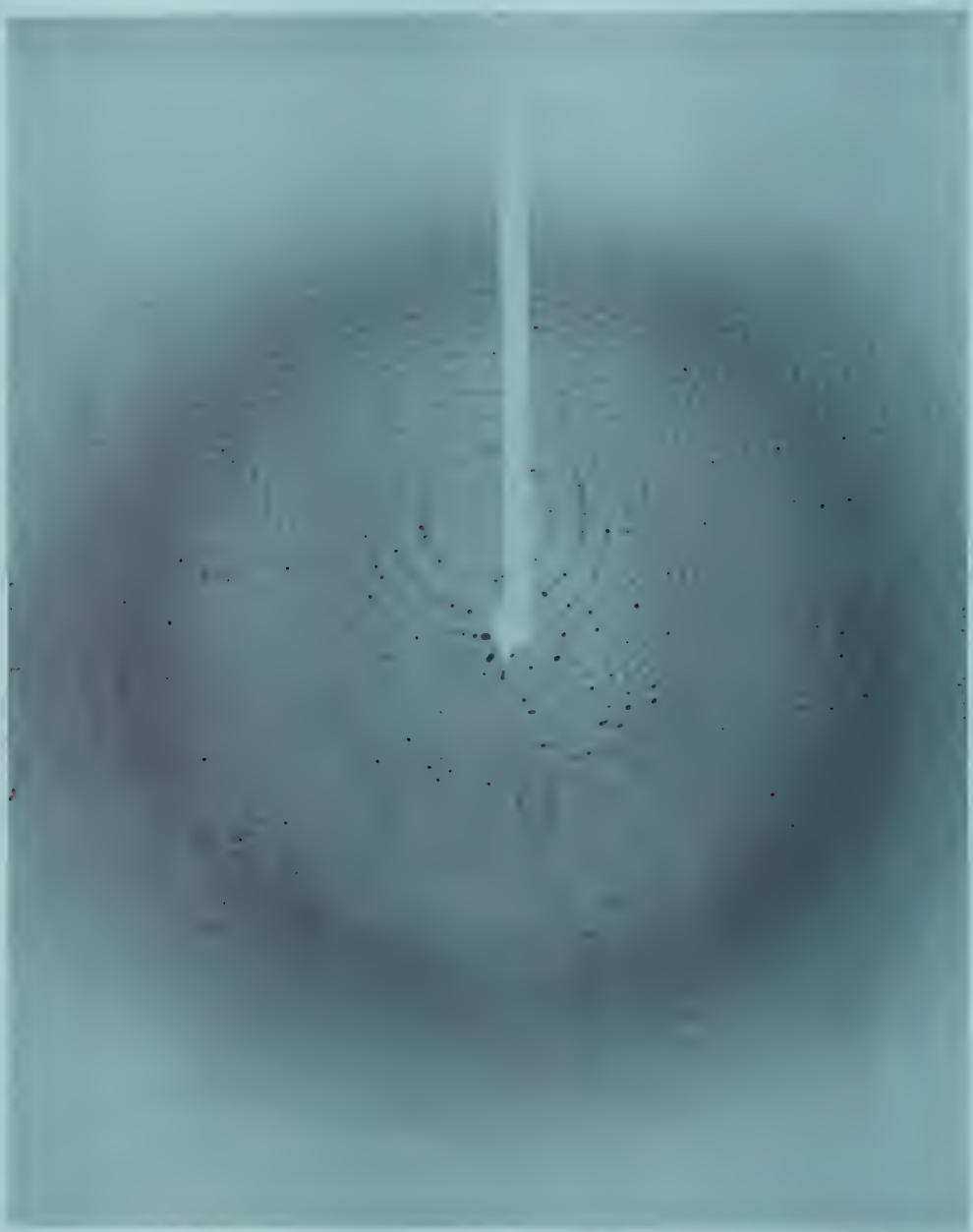


Figure 10. X-ray diffraction photograph of a crystal of poliomyelitis virus. The perfection of the crystal is shown by the presence of reflexions down to spacings of 2 Å. (Photograph by Dr J. T. Finch).

array but with their ends not lined up. X-ray studies of these gels, which have been extensively reviewed (e.g. Crick & Kendrew), have shown that a TMV particle, which is about 3,000 Å long and 150 Å in diameter, is made up of identical (or at least very similar) protein sub-units in a helical arrangement. It has also been discovered that the surface of the virus is grooved or serrated, that it has a hole of radius 20 Å down its centre and that the RNA (the infective component which is present as a long chain) is located, probably in a coil, near a radius of 40 Å. These results, which are due largely to the work of Dr D. L. D. Caspar at Yale and the late Dr Rosalind Franklin and her colleagues at Birkbeck College, London, have been obtained in part by use of the isomorphous replacement technique using particles to which lead or mercury atoms were attached. They illustrate again the great power of the X-ray method when it is applied to suitable material.

As a final example there is the very recent work on the poliomyelitis virus by Finch & Klug of Birkbeck College. Poliomyelitis is a small spherical virus which forms very perfect single crystals, the perfection of which may be judged from the photograph of the diffraction pattern shown in Fig. 10. By analysis of this diffraction pattern Finch & Klug have discovered that the virus particles form a crystal structure with a unit-cell 353 × 378 × 320 Å. The particles themselves are about 300 Å in diameter and they appear to be built up out of sixty identical, or very similar, protein sub-units which form a roughly spherical shell. As in TMV the function of this protein shell probably is to protect the RNA inside but neither the configuration of this RNA nor the way in which the protein shell is made up have yet been determined. Nevertheless it is hardly too much to expect that within the next decade even these complex structures will be completely determined.

REFERENCES

Arndt, U. W. & Riley, D. P. (1955) The structure of some proteins as revealed by an X-ray scattering method. *Phil. Trans. Roy. Soc.* A, **247**, 409.

Bamford, C. H., Elliott, A. & Banby, W. E. (1956) *Synthetic Polypeptides.* Academic Press, New York.

Bluhm, M. M., Bodo, G., Dintzis, H. M. & Kendrew, J. C. (1958) The crystal structure of myoglobin—IV. A Fourier projection of sperm-whale myoglobin by the method of isomorphous replacement. *Proc. Roy. Soc.* A, **246**, 369.

Bragg, W. L. (1939) *The Crystalline State: A General Survey.* Bell, London

Bragg, W. L. & Perutz, M. F. (1954) The structure of haemoglobin—VI. Fourier projections on the 010 plane. *Proc. Roy. Soc.* A, **225**, 315.

Bunn, C. W. (1945) *Chemical Crystallography.* Clarendon Press, Oxford.

Cowan, P. M., North, A. C. T. & Randall, J. T. (1955). *Fibrous proteins and their Biological Significance.* Cambridge Univ. Press.

Crick, F. H. C. & Kendrew, J. C. (1957) X-ray analysis and protein structure. *Advances in Protein Chemistry*, **12**, 133. Academic Press, New York.

Crowfoot, D., Bunn, C. W., Rogers-Low, B. W. & Turner Jones, A. (1949) from *The Chemistry of Penicillin*, p. 310. University Press, Princeton.

Feughelman, M., Langridge, R., Seeds, W. E., Stokes, A. R., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F., Barclay, R. K. & Hamilton, L. D. (1955) Molecular structure of deoxyribose nucleic acid and nucleoprotein. *Nature, London*, **175**, 834.

Finch, J. T. & Klug, A. (1959) Structure of poliomyelitis virus. *Nature, London*, **183**, 1709.

Franklin, R. E. & Gosling, R. G. (1953) Molecular configuration of sodium thymonucleate. *Nature, London*, **171**, 740.

Green, D. W., Ingram, V. M. & Perutz, M. F. (1954) The structure of haemoglobin—IV. Sign determination by the isomorphous replacement method. *Proc. Roy. Soc.* A, **225**, 287.

Hodgkin, D. C. (1958) X-ray analysis and the structure of vitamin B_{12}. *Progress in the Chemistry of Organic Natural Products*, **15**, 167. Springer-Verlag, Vienna.

Hodgkin, D. C., Kamper, J., Lindsey, J., Mackay, M., Pickworth, J., Robertson, J. H., Shoemaker, C. B., White, J. G., Prosen, R. J. & Trueblood, K. N. (1957). The structure of vitamin B_{12}—I. An outline of the crystallographic investigation of vitamin B_{12}. *Proc. Roy. Soc.* A, **242**, 228.

Hodgkin, D. C., Pickworth, J., Robertson, J. H., Prosen, R. J., Sparks, R. A. & Trueblood, K. N. (1959) The structure of vitamin B_{12}—II. The crystal structure of a hexacarboxylic acid obtained by the degradation of vitamin B_{12}. *Proc. Roy. Soc.* A, **251**, 306.

Interatomic Distances (1958) London: The Chemical Society.

James, R. W (1948) *Optical Principles of the Diffration of X-Rays*· Bell: London.

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wykoff, H. & Phillips, D C. (1958) A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature, London*, **181**, 662.

Kendrew, J. C. & Perutz, M. F. (1957) X-ray studies of compounds of biological interest. *Annual Review of Biochemistry*, **26**, 327.

Lindsey, J. M. & Barnes, W. H. (1955) The crystal and molecular structure of codeine hydrobromide dihydrate. *Acta Cryst.*, **8**, 227.

Lipson, H. & Cochran, W. (1953) *The Determination of Crystal Structures*. Bell, London.

Pauling, L., Corey, R. B. & Branson, H. R. (1951) The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U.S.*, **37**, 205.

Perutz, M. F. (1958) Some recent advances in molecular biology. *Endeavour*, **17**, 190.

Porter, A. B. (1906) On the diffraction theory of microscopic vision. *Phil. Mag.*, **11**, 154.

Robertson, J. M. (1953) *Organic Crystals and Molecules*. Cornell University Press, New York

Structure Reports (1940–51) Vols. 8–15. Oosthoek, Utrecht: N.V.A.

Watson, J. D. & Crick, F. H. C. (1953) A structure for deoxyribose nucleic acid. *Nature, London*, **171**, 737. Genetical implications of the structure of deoxyribonucleic acid. *Nature, London*, **171**, 964.

Wilkins, M. F. H., Stokes, A. R. & Wilson, H. R. (1953) Molecular structure of deoxypentose nucleic acids. *Nature, London*, **171**, 738.

FLUORESCENT ANTIBODY TECHNIQUES

R. G. White

The protein chemist, faced with the problem of the specific identification of macromolecules in solution, naturally turns for help to the immunological method. Similarly, an obvious approach to the study of the distribution of a macromolecular substance in tissue sections is by use of specific antibody globulin which has been labelled with a colour or a radioactive tracer. Marrack (1934) clearly showed that this was possible by the demonstration that typhoid bacilli would agglutinate into red clumps by treating a suspension of the organisms with antibody labelled with the red dye R-salt azobenzidine. However the detection of the redness of individual bacteria in the bright field of the microscope was only barely possible. Accordingly Marrack's experiment was repeated with visual advantage by Coons, Creech & Jones (1941) using a fluorescent label anthracene. However the use of this blue fluorescent dye for the detection of an antigen in a tissue section was unsatisfactory since many tissue elements normally fluoresce a bright blue. Accordingly, Coons, Creech, Jones & Berliner (1942) substituted the green fluorescent dye fluorescein, similarly linked to antibody by the carbamido linkage, and such was its success that since 1942 until quite recently this dye has monopolized immune-histology. The reasons for this success are: first, the brilliance of fluorescein which is able to convert its absorbed light energy with about 80 per cent efficiency, secondly, the peak emission at 5,210 Å (yellow-green) which is near the maximum sensitivity of the retina, and thirdly, the rarity of green-fluorescing materials in animal tissues. This latter statement holds only if the viewing system uses light all visible wave-lengths, since many blue or grey natural fluorescent objects emit light containing yellow-green wave-lengths and appear green when viewed with a minus-blue filter.

Coons & Kaplan (1950) described the preparation of amino-fluorescein I and II, from the crude melt produced by heating 4-nitrophthalic acid with two equivalents of resorcinol, by fractional crystallization of the diacetates and subsequent reduction of the separate nitro-fluorescein isomers with hydrogen in the presence of Raney nickel. The purified

nitro-fluorescein (either I or II) was converted to isocyanate by reaction with phosgene in acetone. The isocyanates I and II are unstable, and therefore it was recommended that isocyanate should be prepared from stock fluorescein amine and used immediately. However, as pointed out by Marshall (1951) a solution of fluorescein isocyanate in dry acetone may be stored in the dark at −20° C. without appreciable deterioration for at least one year. The two amino-fluoresceins I and II (both are satisfactory as labels) were analysed chromatographically by De Repentigny & James (1954) and found to include 5–10 per cent impurities. Their chromatographic method provides an alternative preparative procedure for the amino-fluoresceins.

In the method of Coons & Kaplan (1950) for the preparation of isocyanate, phosgene, obtained as a gas from a cylinder, was bubbled through an acetone solution of the fluorescein amine. This is an undertaking which cannot be accepted light-heartedly especially in a hospital laboratory. As an alternative, ampoules of a solution (13 per cent by weight approximately) of phosgene in toluene, or preferably benzene, may be used (British Drug Houses, Ltd). Fluorescein amine, dissolved in acetone is dropped slowly from a separating funnel into the benzene solution of phosgene. After allowing the reaction mixture to stand for 30 minutes with occasional shaking, the acetone and benzene can be distilled off under vacuum at 45° C. The author has found this procedure (worked out in collaboration with Dr H. Gooder) a very convenient and reliable substitute for the classical procedures. Riggs, Seiwald, Burckhalter, Downs & Metcalf (1958) have provided a further simplification by the synthesis of stable, solid isothiocyanate of fluorescein, which has, in the author's experience proved to be a very convenient alternative to fluorescein isocyanate.

The coupling of isocyanate to protein is carried out quite simply (Coons, 1958) by adding fluorescein amine dissolved in acetone-dioxane mixture slowly to the ice-cold protein solution in physiological saline buffered to pH 9·0 in the presence of dioxane and acetone. Coons & Kaplan (1950) found the optimum ratio of protein to fluorescein amine to be 20 : 1. When the isocyanate was increased the number of fluorescein-carbamido groups introduced remained constant, indicating under these conditions a maximum of about two groups per molecule of protein (rabbit type III anti-pneumococcus). Baldwin *et al.* (1959) have examined in detail the chemical and immunochemical properties of conjugates prepared with the isocyanate of 4-dimethyl amino azo-benzene (D.A.B.). It is of interest that whereas a maximum of sixty-five D.A.B. groups

could be introduced into bovine serum albumin, it was not possible with bovine γ-globulin to introduce more than three groups per molecule without resulting loss of solubility and denaturation of the product.

Goldman & Carver (1957) have further simplified the conjugation procedure by using fluorescein isocyanate dried on to thick filter paper. The addition of an appropriately sized piece to stirred, buffered antibody solution allowed conjugation without use of any organic solvent.

Two or more colours would be a great advantage in many fluorescence studies. Clayton (1954) described the use of nuclear-fast red, a dye which combined firmly with protein producing crimson fluorescence. Unfortunately the chemical nature of the dye was not known, although it is apparently not a rhodamine.

Clayton (1954) also described the use of 1-dimethyl amino naphthalene-5-sulphonyl chloride, a lemon-yellow fluorochrome developed by Weber (1952) which has been more extensively applied by Redetzki (1958) and Mayersbach (1958). Details of the synthesis, the method of conjugation of the dye to protein and the immunological properties of the conjugate are recorded in the papers by Laurence (1957), and Mayersbach (1958). Also, recently two rhodamines have been described, whose red fluorescence provides excellent contrast to fluorescein isocyanate in a doubly-stained section (Silverstein, 1957; and Chadwick, McEntegart & Nairn, 1958a and b). The latter authors used Lissamine rhodamine RB 200 which was linked rapidly and conveniently to protein through a sulphonamido group.

Both fluorescein and rhodamine protein conjugates are satisfactorily stable over a year or longer and even when in frequent use are not often subject to contaminating growth of bacteria or fungi. Merthiolate (1 : 10,000) is usually added as a preservative but, in my experience, capsulated *Bacterium aerogenes* occasionally grows in spite of this and has necessitated the addition of sodium azide (0·3 per cent final concentration).

In the preparation of tissue sections for use with fluorescent antibody solutions the procedures must always avoid damage to the immunological specificity of the antigen and, at the same time prevent the movement of antigen by solution or otherwise. Polysaccharide antigens in general might be expected to withstand treatment with the standard histological fixatives, and the organic solvents required for embedding the tissue in paraffin wax. Coons & Kaplan (1950) and Hill, Deane & Coons (1950) used mouse tissues fixed in cold picric acid-alcohol-formalin and subsequently embedded in paraffin for the localization of polysaccharides of *Streptococcus pneumoniae* and Friedländer's bacillus.

Proteins have been successfully processed by the cutting of unfixed frozen tissue blocks in the Linderstrøm-Lang cryostat at −16° C. to −18° C., yielding sections of 3–5 μ thickness. The use of the knife attachment described by Coons, Leduc & Kaplan (1951) greatly facilitates this operation. The freeze-drying preparation of tissue has also been used after rapid quenching of small blocks of tissue in liquid nitrogen-isopentane mixture. With this procedure ice-crystal artifacts should be minimal. When frozen sections are obtained in a cryostat, presumably the act of cutting involves thawing and re-freezing at the knife cutting edge and the opportunity for extensive ice crystal artifact. Another advantage of the freeze-drying method is that 2 μ sections may be cut, allowing sharper definition of the subsequent fluorescent image. In general, it would seem good tactics when attempting the localization of a new antigen to use frozen sections prepared in a cryostat and to reserve the freeze-drying or freeze substitution (Bell, 1956; Feder & Sidman, 1958) methods for a more refined histological approach if this proves necessary.

The frozen section from the cryostat, thawed on to a clean slide and dried rapidly in a current of warm air, next requires treatment with a suitable fixative. For proteins the following have been used: 95 per cent v/v ethanol at 37° C. (Coons *et al.*, 1951), absolute methanol at 0° C. (Gitlin, Landing & Whipple, 1953), formalin-dioxane (Marshall, 1954). Certain antigens e.g. of basement membranes (Hill & Cruickshank, 1953) and of *Paramoecium* (Beale & Kacser, 1957) have been left unfixed. For viruses, acetone has been found generally successful. When virus containing cells have been left unfixed (O'Dea & Dineen, 1957) a fluorescence limited to the cell periphery resulted presumably since the living cytoplasmic membrane was impermeable to globulins.

The use of fluorescent globulin conjugates must be preceded by manoeuvres to avoid certain non-specific interactions with sundry tissue elements. Fig. 3 gives such an example of non-specific fluorescence in this case to neutrophil leucocytes. Non-specific staining may result for various reasons. Some may result from antibody to tissue antigens. Thus rabbit antisera could contain Forssman antibody which after conjugation would combine with the erythrocytes and other tissue structures (Tanaka & Leduc, 1956) in a Forssman-positive species like the sheep. Secondly, some could be due to unreacted fluorescent compound adsorbed on to the proteins of the conjugate and not removed by dialysis. Thirdly, all normal serum contains proteins which react with tissue components (Kidd & Friedewald, 1942; Berenbaum, 1958). The most widely-used method for removing such unwanted reactions is by absorption of the conjugate with

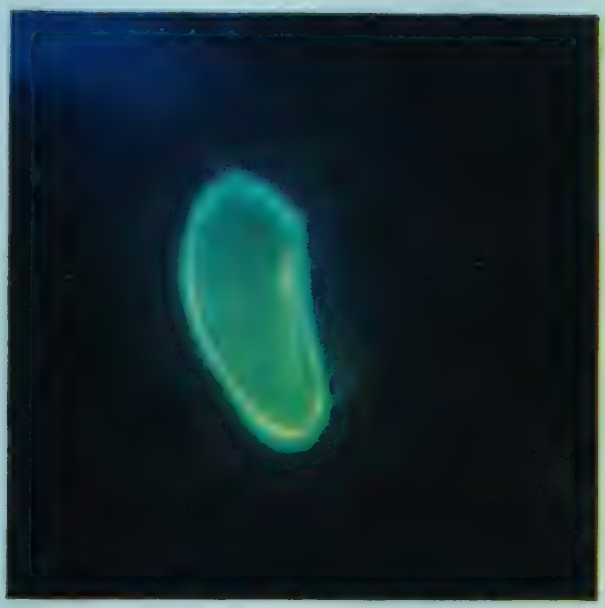


Figure 1. *Paramecium aurelia.* A conjugating pair of paramecia of different antigenic types. After conjugation had proceeded for some hours and the mates were firmly joined together (along the flattened lower side of the organism shown), the organisms were fixed with osmic acid, and treated with fluorescent antibody effective against one but not the other member of the pair. The resultant precise demonstration of the homologous organism only bears witness to the complete specificity of this result.

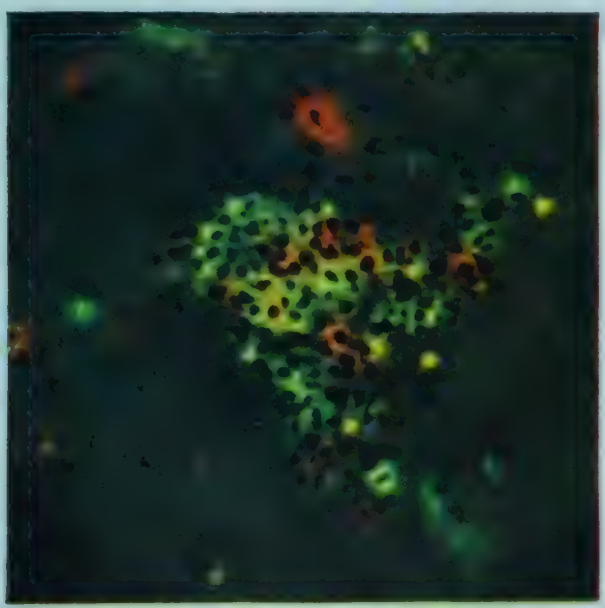


Figure 2. Antibody-containing cells in lymph node. Plasma cells of varying grades of maturity containing diphtheria antitoxin (rhodamine, red) and anti-ovalbumin (fluorescein, green) in the popliteal node of a rabbit four days after secondary response to both antigens injected into the footpad. Sandwich technique.

tissue powder suspensions e.g. acetone dried mouse liver powder (Coons & Kaplan, 1950). Even repeated absorptions with such powder may fail to remove all non-specific staining e.g. of polymorph neutrophil myelocytes and leucocytes. However, such residual staining can sometimes be dealt with by specific treatment e.g. in the case of the polymorphs by absorption with bone marrow powder (Sheldon, 1953) or by the use of the anionic exchange resin Dowex-2 chloride (Coons, 1958). Chadwick *et al.* (1959) have estimated that after repeated absorptions of conjugates with tissue powder 25–50 per cent of the unreacted fluorescent compound

Figure 3. Non-specific fluorescence of neutrophil myelocytes and leucocytes in frozen section of lymph node treated with a fluorescein conjugate.

still remains adsorbed. This remainder can be removed by shaking one hour with activated charcoal, at sacrifice of some potency but with gain in specificity of staining. Another method of dealing with non-specific staining is by contrast staining. Thus if, say, a fluorescein conjugate has been used to render a specific antigen green, subsequent use of a rhodamine conjugate of normal serum on the same section will convert most non-specifically stained structures to a contrasting orange or red.

Controls for establishing the specificity of staining are always necessary. In the case of an antigen foreign to the tissue used, the normal tissue from

the same species is an obvious primary control. Further control can be effected by the blocking effect of prior or simultaneous treatment of the section with unconjugated specific antibody. Also, the conjugate can itself be absorbed with antigen. Even when the antibody of the fluorescent conjugate is non-precipitating, specific inhibition can be achieved by addition of excess of purified antigen to the conjugate.

Controls for the staining of a normal tissue constituent are more difficult, although where something of the antigen distribution is known the problem is simplified. Thus Tanaka & Leduc (1956) could control the staining of Forssman antigen by the use of analogous tissue from non-Forssman containing species. In other cases e.g. the basement membrane

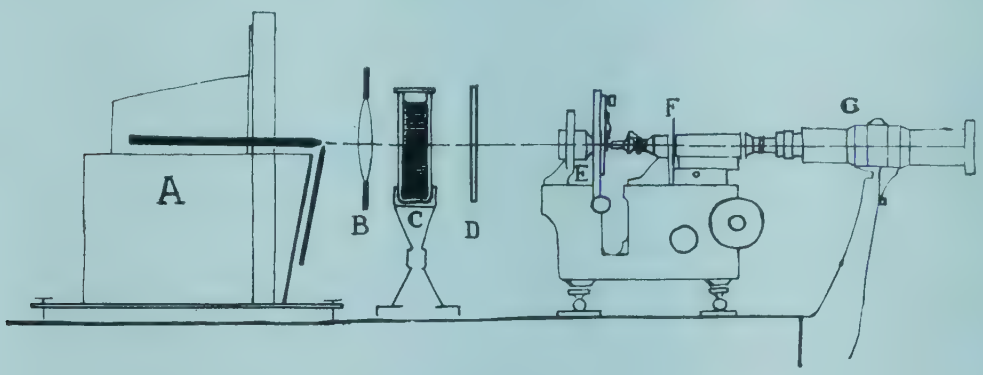


Figure 4. Diagram of fluorescence microscope.
A. Carbon arc.
B. Condensing lens.
C. Cuvette of 20 per cent copper sulphate solution.
D. Filter.
E. Bi-reflecting type dark field condenser.
F. Filter.
G. Camera.

antigen of the renal glomerulus (Hill & Cruickshank, 1953), or A.C.T.H. in the anterior pituitary (Marshall, 1951) the problem of establishing specificity of staining is complex and individual, since the antigens used for preparation of the conjugate are usually impure (see further discussion below).

For fluorescence microscopy an intense source of long wave ultra-violet radiation is essential. In spite of its inconveniences of heat and noise the carbon arc (Fig. 4, A) on direct current provides a circular source of beautifully even intensity at the positive crater. With a glass or quartz condensing lens (B) of accurately worked aspherical contours the image of this source can be adjusted to fill exactly the back lens of the condenser (E). Otherwise the more generally convenient high pressure mercury

vapour arc of 250 watts or 1 kW. may be used. The filters (D and F) used will vary with the absorption and emission spectra of the fluorescence conjugates chosen, and details may be obtained in the papers relating to the use of individual fluorochromes. The condenser (E) is a bi-reflecting type dark field condenser of glass.

The sensitivity of the method as judged by the lowest amount of a biological substance capable of detection is impressively high. Thus the O antigen in a single smooth *Shigella dysenteriae* bacillus can be easily detected. Morgan (1949) has estimated this antigen at 5–7 per cent of the dry bacillary weight, and a single bacillus would thus contain approximately $2{\cdot}5 \times 10^{-12}$ mg. However, the histological detection of any antigen will demand that this be present in high biological concentration and in a situation which gives good contrasting background. Soluble protein antigens can be readily detected in the blood within sectioned vessels at a concentration of 200 μg./ml., and antibody (by the sandwich technique) at 100 μg./ml.

There exist several modifications of the technique which use two or more immunologically-linked layers for the demonstration of antigen or antibody. The so-called sandwich technique employs a primary layer of a dilute solution of unconjugated antigen. After reacting for 30 minutes, this is rinsed off with buffered saline, washed for 10 minutes and then exposed to specific fluorescent antibody (Coons, Leduc & Connolly, 1955). Other double-layer techniques to detect antigen have used labelled anti-globulin sera as a second layer in order to detect the sites of localization of an unlabelled specific antibody which was applied as the first layer. The sensitivity of the method is increased by such double-layer techniques since each added layer combines as an antibody with subjacent multivalent antigen. The method can be successful even though the antibody in the primary layer is much diluted. Recently, Goldwasser & Shepard (1958) have shown that guinea-pig complement can be used in the staining of antigens in combination with specific unlabelled antisera. It is necessary to remember that absorption of the unlabelled sera used in such multi-layer techniques will usually be necessary in order to avoid non-specific reactions.

The photographic reproduction of the results of fluorescent-antibody studies is difficult since full interpretation of the image often depends on colour. The black and white reproductions do not enable a distinction between the specific fluorescence of the fluorochrome and the natural fluorescence of the tissue. The alternative of colour photography is also unsatisfactory on account of the imperfect colour rendering. Images of

the same colour but different brightness will appear to be of different colour in the coloured reproduction. Thus in Fig. 1 the intensely green fluorescence of part of the specimen is reproduced a yellow colour. The antibody-containing cells of Fig. 2 also appear to be of several different shades of green (or red), whereas in the original preparation they were all of the same colour of either red or green. Any quantitative reproduction of the data is beset with enormous difficulties, although a preliminary attempt by Mellors, Siegel & Pressman (1955) employed a yellow filter and relative measurements of the fluorescence intensity of the image was made by photographic photometry.

Some Applications of the Technique

The potentialities of labelled antibody as a histochemical tool are considerable. In point of achieved fact, the past five or seven years have seen the application of fluorescent antibody to so many fields of endeavour

Figure 5. Capsular polysaccharide antigen of type III pneumococcus in mouse kidney. The bright areas represent the antigen in the loose connective tissue between the dark circles of the cortical tubules. Fluorescein conjugate of anti-polysaccharide antibody. Single-layer technique.

that it is impracticable, in this paper, to mention more than a small fraction of the published works.

The first essays involved the tracing of injected foreign antigens, and

the localization in their mammalian hosts of viruses and other pathogenic agents. These are, of course, by far the easiest problems to tackle since, first, the antigens were already well characterized immunologically, and secondly, controls for the specificity of the staining results were relatively easy to apply. These controls are necessary to show that the fluorescent image is neither an example of natural fluorescence, nor the result of the non-specific absorption of any of the mixed bag of labelled proteins which accompany the specific antibody in the conjugate, nor are unexpected cross-reactions with a normal tissue constituent. With a foreign antigen it is easy to comply with these since normal tissue is available for comparison

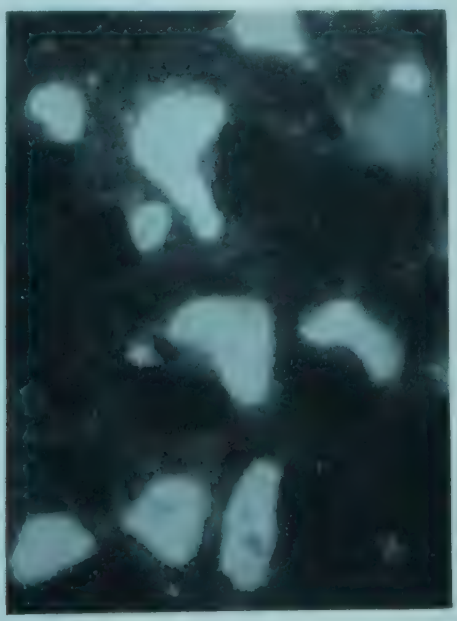


Figure 6. Lymphogranuloma venereum virus in HeLa cells. The bright areas are cytoplasmic inclusion bodies. The dark areas are nuclei of the tissue culture cells. Twenty-four hours after inoculation. Double-layer technique with human antiserum and fluorescein conjugate of rabbit anti-human gamma-globulin.

and the reaction can often be specifically blocked by antigen or antibody prepared without relation to tissue components.

Fig. 5 shows the pattern of localization of the capsular polysaccharide of *Str. pneumoniae* in a paraffin section of the cortex of mouse kidney. Kaplan, Coons & Deane (1950) and Hill, Deane & Coons (1950) described the distribution of several of such injected polysaccharide macromolecules (m.w. 100,000–500,000) which included synovial cells, osteoblasts and the cells of the distal convoluted kidney tubules, in addition to reticulo-

endothelial cells. The macrophages showed a content of these antigens for six months at least and presumably incapacity of these cells to destroy the antigen forms the basis of the remarkable phenomenon of immunological paralysis (Felton, 1949). In later papers Coons, Leduc & Kaplan (1951) using cryostat-prepared frozen sections studied the distribution of injected heterologous proteins and Gitlin, Landing & Whipple (1953) applied the same method to study homologous protein in man. These interesting findings and their comparison with the results from using injected fluorescein and radioactive labelled proteins are summarized by Coons (1956).

Coons, Snyder, Cheever & Murray (1950) early showed that mumps

Figure 7. Proteus vulgaris in Kupffer cells of mouse liver. Discrete bacillary outlines, as well as free antigen are seen as bright fluorescence. Single-layer technique with fluorescein conjugate of rabbit antibody to the O (somatic) antigen.

virus could be microscopically localized in sections of the parotid of experimentally infected monkeys, and since then an impressive list of successes with a variety of virus agents have been recorded. The visualized material may consist either of single or aggregated virus particles or antigenic material derived from these (Liu, 1955). Fig. 6 shows as bright circumscribed areas the cytoplasmic inclusion bodies of the virus of lymphogranuloma venereum growing within HeLa cells in monolayer

tissue culture. Bacteria can be revealed by the use of fluorescent antibody directed against their surface agents. Fig. 7 shows *Proteus* organisms within Kupffer cells of a frozen section of mouse liver (Wood & White, 1956). The rapid and specific identification of micro-organisms has many possible applications, e.g. in the characterization of *Salmonella* species in faeces, *Streptococcus pyogenes* in the throat and *Neisseria gonorrhoeae* in urethral smears. Further, a standard culture of organisms may be used to identify or titrate antibody in sera as a diagnostic procedure. Thus, Deacon, Falcone & Harris (1957) have proposed a slide test using patient's serum and fluorescent anti-human globulin reagent for the detection of antibody to *Treponema pallidum* as a rapid and convenient alternative for the treponema immobilization test.

Another major field in which a modification of the technique has provided definitive results is the localization of specific antibody by the so-called 'sandwich' technique (Coons, Leduc & Connolly, 1955; Leduc Coons & Connolly, 1955; White, Connolly & Coons, 1955a and b; White, 1954; see also Fig. 2). In a variety of situations (lymph nodes, spleen, subcutaneous granuloma) the cells containing antibody were shown to correspond to members of a differentiating family of cells whose mature member was the plasma cell. Antibody was most often seen as a diffuse fluorescence outlining the cytoplasm and in localized intra-nuclear spots. Sometimes the cytoplasmic antibody was in the form of very bright spherical bodies. These apparently correspond to the eosinophilic clusters of spherical bodies within plasma cells which are quite common in chronic granulomatous lesions (Russell bodies) and were originally regarded as a fungus.

Finally, the distribution of several antigenic substances which are native to animal tissues has been explored, e.g. the antigenic constituents of kidney glomerular basement membrane which can induce the formation of Masugi nephrotoxin, the hormones of the anterior pituitary, pancreatic enzymes, thyroid auto-antigens, gamma-globulin and other serum proteins, blood group substances, lens substance, muscle myosin and liver antigens. The difficulties involved are well illustrated by the attempt to localize adrenocorticotrophic hormone. Marshall (1951) prepared antiserum in adrenalectomized rabbits against A.C.T.H. prepared commercially from swine pituitaries. A fluorescein conjugate of the globulin fraction, secured by ethanol fractionation of the rabbit antiserum, was applied to freeze-dried sections of hog, sheep and beef pituitaries. This crude conjugate, as might be expected, stained plasma proteins within the blood vessels but this effect could be removed by treating the conjugate

with hog serum. This left a brilliant staining reaction against the basophil cells of hog pituitary and no reaction against the pituitaries of other species. The nuclei did not show fluorescence but the cytoplasm appeared full of densely-packed granules. This result would appear to fit with the other indirect evidence which attributes the secretion of A.C.T.H. to the basophils, i.e. the findings by Crooke & Russell that hyaline changes occur in these cells in Cushing's syndrome and in Addison's disease. However, Marshall's paper contained no evidence that the antiserum which he used was really hormone specific, i.e. would combine only with purified A.C.T.H. and not with purified preparations of thyrotropic hormone or growth hormone. In a subsequent communication Cruickshank & Currie (1958) found that rabbit antisera to A.C.T.H. always cross-reacted with T.S.H. and growth hormone preparations. These results certainly make it clear that studies of normal tissue components require the support of comprehensive preliminary serological testing.

The preparation of the appropriate fluorescent antibody conjugate against a normal tissue constituent is seldom a straightforward exercise. Given the antigenicity of the tissue component, its purification is likely to afford the greatest problem, and in this connection it should be realized that small amounts of contaminating antigens in the material used for stimulating antibody production in another animal may result in disproportionate amounts of contaminating antibody. However, occasionally nature is herself very helpful in providing the antibody, and in human cases of Hashimoto's disease, the serum may contain 5 mg./ml. of autoantibody against thyroglobulin, which compares very well with the level of specific antibody in the serum of a hyper-immune rabbit. The result of applying the fluorescein conjugate of the globulin of such a serum to a frozen section of the patient's own thyroid is shown in Fig. 8. Similarly, the sera of systemic lupus cases will localize to nuclei of human and animal tissues, since they apparently contain antibody to a normal nuclear constituent, possibly nucleo-histone. In Figs. 9, 10 and 11 are compared the results of applying to a frozen section of thyroidectomy specimen of a gland of more or less normal histology, the sera from a case of Hashimoto's disease (localization to the intra-acinar colloid), from a case of systemic lupus (localization to the nuclei), and from a case showing features of both diseases (localization to both nuclei and colloid).

ACKNOWLEDGMENTS

The author will always owe a great personal debt to Dr A. H. Coons, from whose writings or personal tuition has been derived much of the

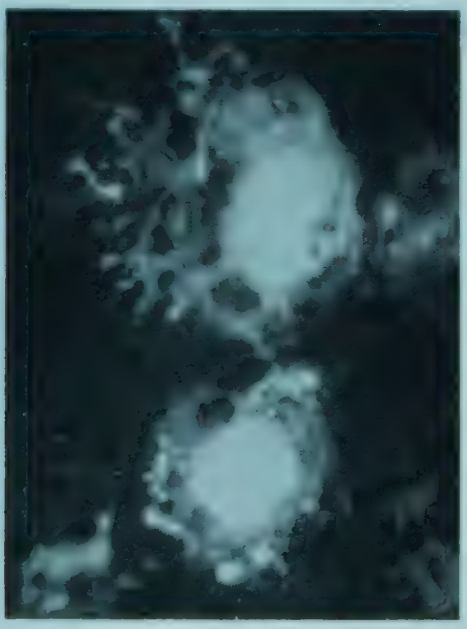


Figure 8. Thyroid gland from case of Hashimoto's disease. Two small acini contain auto-antigen (thyrogobulin) which is present outside the acini in and among the cells of the granulomatous infiltration. Single-layer technique employing fluorescein conjugate of the serum globulin on a frozen section of the thyroid of the same patient.

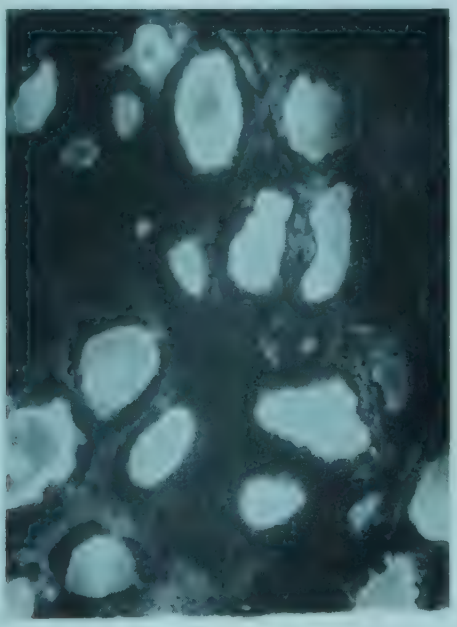


Figure 9. Antibody analysis of human sera. Thyroglobulin antigen in a normal thyroid, revealed as bright intra-acinar areas have localized antibody from the serum of a patient of Hashimoto's disease. Double-layer technique employing serum of case of Hashimoto's disease and a fluorescein conjugate of rabbit anti-human gamma-globulin.

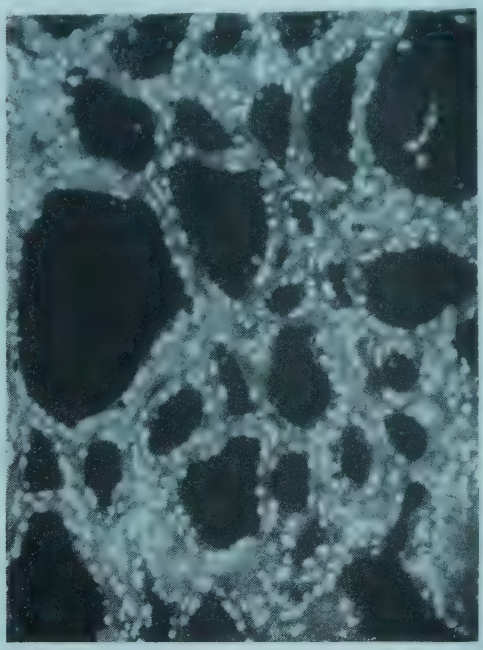


Figure 10. Antibody analysis of human sera. Nuclear antigen in a normal thyroid, revealed as bright areas (nuclei) which have localized the antibody from the serum of a human case of systemic lupus erythematosus. Double-layer technique employing serum of case of systemic lupus and a fluorescein conjugate of rabbit anti-human gamma-globulin.

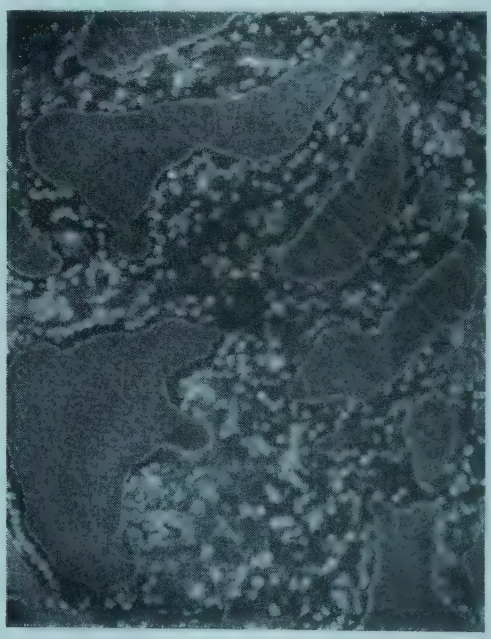


Figure 11. Antibody analysis of human sera. This result, which is a combination of those shown in Figs. 9 and 10 was given by exposing a section of normal thyroid tissue to the serum of a patient suffering from both systemic lupus erythematosus and Hashimoto's disease. Double-layer technique employing serum of patient and a fluorescein conjugate of rabbit anti-human gamma-globulin.

experience on which this lecture was based. I am very grateful to Dr G. H. Beale for permission to include the colour photograph of *Paramecium aurelia*.

REFERENCES

Baldwin, R. W., Beswick, J. E., Chayen, J. & Cunningham, G. J. (1959) Immunochemica studies of protein-carcinogen binding in rat liver during azo dye carcinogenesis. *Acta Unio. contra Cancrum* (in Press).

Beale, G. H. & Kaeser, H. (1957) Studies on the antigens of *Paramecium aurelia* with the aid of fluorescent antibodies. *J. gen. Microbiol.*, **17**, 68–74.

Bell, L. G. E. (1956) *in Physical techniques in Biological Research*, **3**, 1–27. Academic Press, New York.

Berenbaum, M. C. (1959) The autoradiographic localization of intracellular antibody. *Immunology*, **2**, 71–83.

Chadwick, C. S., McEntegart, M. G. & Nairn, R. C. (1958a) Fluorescent protein tracers. A simple alternative to fluorescein. *Lancet*, i, 412–414.

Chadwick, C. S., McEntegart, M. G. & Nairn, R. C. (1958b) Fluorescent protein tracers. A trial of new fluorochromes and the development of an alternative to fluorescein. *Immunology*, **1**, 315–327.

Chadwick, C. S., Nairn, R. C. & McEntegart, M. G. (1959) Investigation of the unreacted fluorescent material in fluorescein-protein conjugates. *Biochem. J.*, **73**, 41P.

Clayton, Ruth, M. (1954) Localization of embryonic antigens by antisera labelled with fluorescent dyes. *Nature, Lond.*, **174**, 1059.

Coons, A. H., Creech, H. J. & Jones, R. N. (1941) Immunological properties of an antibody containing a fluorescent group. *Proc. Soc. exp. Biol. Med.*, **47**, 200–202.

Coons, A. H., Creech, H. J., Jones, R. N. & Berliner, E. (1942) The demonstration of pneumococcal antigen in tissues by the use of fluorescent antibody. *J. Immunol.*, **45**, 159–170.

Coons, A. H. & Kaplan, M. H. (1950) Localization of antigen in tissue cells.—II. Improvements in a method for the detection of antigen by means of fluorescent antibody. *J. exp. Med.*, **91**, 1–13.

Coons, A. H., Snyder, J. C., Cheever, F. S. & Murray, E. S. (1950) Localization of antigen in tissue cells.—IV. Antigens of Rickettsiae and Mumps Virus. *J. exp. Med.*, **91**, 31–38.

Coons, A. H., Leduc, E. H. & Kaplan, M. H. (1951) Localization of antigen in tissue cells. —VI. The fate of injected foreign proteins in the mouse. *J. exp. Med.*, **93**, 173–188.

Coon, A. H., Leduc, E. H. & Connolly, Jeanne M. (1955) Studies on antibody production. —I. A method for the histochemical demonstration of specific antibody and its application to a study of the hyperimmune rabbit. *J. exp. Med.*, **102**, 49–60.

Coons, A. H. (1956) Histochemistry with labelled antibody. *Int. Rev. Cytol.*, **5**, 1–43.

Coons, A. H. (1958) Fluorescent antibody methods *in General Cytochem. Methods*, ed. J. F. Danielli, **1**, 399–422. Academic Press, New York.

Cruickshank, B. & Currie, A. R. G. (1958) Localization of tissue antigens with the fluorescent antibody technique: application to human anterior pituitary hormones. *Immunology*, **1**, 13–27.

Deacon, W. E., Falcone, V. H. & Harris, A. D. (1957) A fluorescent test for treponemal antibodies. *Proc. soc. exp. Biol. Med.*, **96**, 477–480.

De Rethigny, J. & James, A. T. (1954) A chromatographic separation of the aminofluoresceín isomers. *Nature, Lond.*, **174**, 927–928.

Fisher, N. & Sitman, R. L. (1958) Methods and principles of fixation by freeze-substitution. *J. biophys. biochem. Cytol.*, **4**, 593–602.

Felton, L. D. (1949) The significance of antigen in animal tissues. *J. Immunol.*, **61**, 107–117.

Gitlin, D., Landing, B. H. & Whipple, A. (1953) The localization of homologous plasma proteins in the tissues of young human beings as demonstrated with fluorescent antibodies. *J. exp. Med.*, **97**, 163–176.

Goldman, M. & Carver, R. K. (1957) Preserving fluorescein isocyanate for simplified preparation of fluorescent antibody. *Science*, **126**, 839–840.

Goldwasser, R. A. & Shepard, C. C. (1958) Staining of complement and modifications of fluorescent antibody procedures. *J. Immunol.*, **80**, 122–131.

HILL, A. G. S. & CRUICKSHANK, B. (1953) A study of antigenic components of kidney tissue. *Brit. J. exp. Path.*, **34**, 27–34.

HILL, A. G. S., DEANE, HELEN W. & COONS A. H. (1950) Localization of antigen in tissue. cells.—V. Capsular polysaccharide of Friedländer Bacillus, Type B, in the mouse. *J. exp. Med.*, **92**, 35–44.

KAPLAN, M. H., COONS, A. H. & DEANE, HELEN W. (1950) Localization of antigen in tissue cells.—III. Cellular distribution of pneumococcal polysaccharides types II and III in the mouse. *J. exp. Med.*, **91**, 15–30.

KIDD, J. G. & FRIEDEWALD, W. F. (1942) A natural antibody that reacts *in vitro* with a sedimentable constituent of normal tissue cells.—II. Specificity of the phenomenon: general discussion. *J. exp. Med.*, **76**, 557–578.

LAURENCE, D. J. R. (1957) Fluorescence techniques for the enzymologist *in Methods of Enzymology*, **4**, 174–212. Academic Press, New York.

LIU, C. (1955) Studies on influenza infection in ferrets by means of fluorescein-labelled antibody—II. The role of 'soluble antigen' in nuclear fluorescence and cross reactions. *J. exp. Med.*, **101**, 677–686.

MARRACK, J. R. (1934) Nature of antibodies. *Nature, Lond.* **133**, 292–293.

MARSHALL, J. M. Jr. (1951) Localization of adrenocorticotropic hormone by histochemical and immunochemical methods. *J. exp. Med.*, **94**, 21–29.

MARSHALL, J. M. Jr. (1954) Distribution of chymotrypsinogen, procarboxypeptidase, desoxyribonuclease and ribonuclease in bovine pancreas. *Exptl. Cell Research*, **6**, 240–242.

MAYERSBACH, H. (1958) Immunhistologische Methoden II. Ein weiterer Markierungsfarbstoff: Dimethyl-1-Naphthylaminsulfonsäure–5. *Acta Histochemica*, **5**, 351–367.

MELLORS, R. C., SIEGEL, M & PRESSMAN, D. (1955) *Lab. invest.*, **4**, 69.

MORGAN, W. T. J (1949) *in The Nature of the Bacterial Surface*. Blackwell, Oxford, p. 13.

O'DEA, J. F. & DINEEN, J. K. (1957) Fluorescent antibody studies with *Herpes simplex* virus in unfixed preparations of trypsinized tissue cultures. *J. gen. Microbiol.*, **17**, 19–24.

REDETZKI, H. M. (1958) Labelling of antibodies by 5-Dimethylamino-1-Naphthalene, Sulfonyl chloride. Its effect on antigen-antibody reaction. *Proc. soc. exp. Biol. Med.*, **98**, 120–122.

RIGGS, J. L., SEIWALD, R. J., BURCKHALTER, J. H., DOWNS, C. M. & METCALF, T. G. (1958) Isothiocyanate compounds as fluorescent labelling agents for immune serum. *Am. J. Path.*, **34**, 1081–1092.

SHELDON, W. H. (1953) Leptospiral antigen demonstrated by the fluorescent antibody technique in human muscle lesions of *Leptospirosis icterohaemorrhagiae*. *Proc. soc. exp. Biol. Med.*, **84**, 165–167.

SILVERSTEIN, A. M. (1957) Contrasting fluorescent labels for two antibodies. *J. Histochem. Cytochem.*, **5**, 94–99.

TANAKA, N. & LEDUC, ELIZABETH H. (1956) A study of the cellular distribution of Forssman antigen in various species. *J. Immunol.*, **77**, 198–212.

WEBER, G. (1952) Polarization of the fluorescence of macromolecules.—2. Fluorescent conjugates of ovalbumin and bovine serum albumin. *Biochem. J.*, **51**, 155–167.

WHITE, R. G. (1954) Observations on the formation and nature of Russell bodies. *Brit. J. exp. Path.*, **35**, 365–376.

WHITE, R. G., COONS, A. H. & CONNOLLY, JEANNE M. (1955a) Studies on antibody production.—III. The alum granuloma. *J. exp. Med.*, **102**, 73–82.

WHITE, R. G., COONS, A.H. & CONNOLLY, JEANNE M. (1955b) Studies on antibody production. —IV. The role of a wax fraction of mycobacterium tuberculosis in adjuvant emulsions on the production of antibody to egg albumin. *J. exp. Med.*, **102**, 83–104.

WHITE, R. G. (1957) An immunological investigation of Hashimoto's disease. *Proc. Roy. Soc. Med.*, **50**, 953–954.

WOOD, C. & WHITE, R. G. (1956) Experimental glomerulo-nephritis produced in mice by subcutaneous injections of heat-killed *Proteus vulgaris*. *Brit. J. exp. Path.*, **37**, 49–60.

PHONOCARDIOGRAPHY

AUBREY LEATHAM

Physical Principles and Design of Apparatus for Phonocardiography

Phonocardiography is the graphic registration of vibrations resulting from contraction and dilatation of the heart. If these vibrations are recorded exactly as they arrive at the chest wall there is a tremendous dominance of low frequency vibrations which have so far yielded little information; these vibrations are so slow in movement that it is unlikely that they will ever give valuable information about rapid movements of the heart or of its valves. The high frequency components of cardiac vibrations show a much closer relation to cardiac haemodynamics, and also correspond more closely to the findings of auscultation, for only these high frequency vibrations can be appreciated by the human hearing mechanism. The high frequency components of cardiac vibrations, however, are so small in amplitude in relation to the low that it is unsatisfactory to show both on the same recording (Fig. 1). Severe filtration of the low frequencies, and tremendous amplification of the high, are therefore required to produce the high frequency phonocardiogram which can be closely correlated with cardiac haemodynamics and corresponds to auscultation.

To achieve true recordings without disturbance of the base-line, careful attention must be paid to the design of the apparatus. Extraneous noise and inherent noise in the amplifier are likely to be encountered owing to the great amplification which is required, and the galvanometers must be fully sensitive at 1,000 cycles or more. A sensitive microphone is sealed to the chest wall, but must not be held too tightly, for this stretches the skin into a rigid attenuating diaphragm; a light crystal microphone held on by suction has been the best system in our experience. The amplifier must be specially designed to be free of noise. Filters should cut the low frequencies in the same way as the human hearing mechanism and stethoscope (Fig. 2, the resulting phonocardiogram may be called 'high frequency') in order to achieve standardization, and even more severe

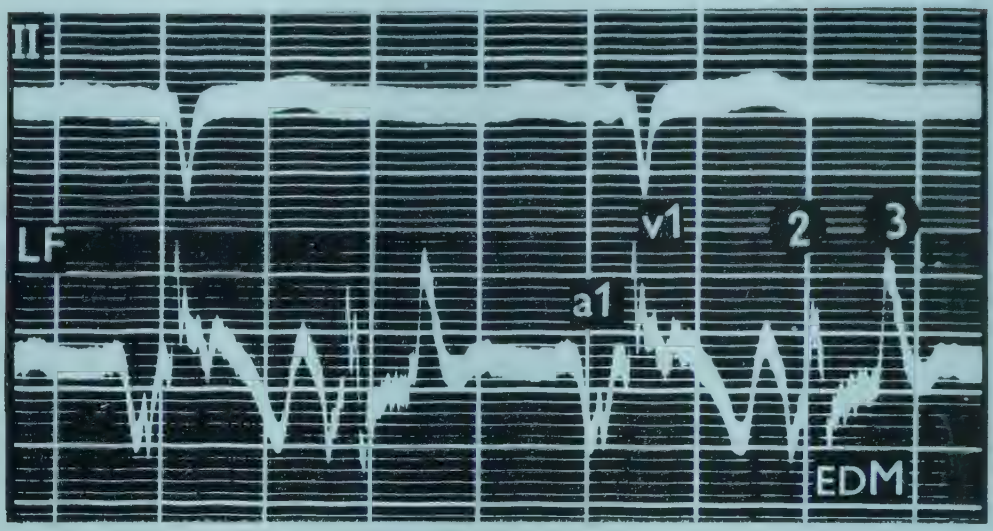


Figure 1. An ultra low frequency phonocardiogram taken during pioneering efforts by Dr William Evans. Some high frequency murmurs can be seen superimposed on the large low frequency oscillations.

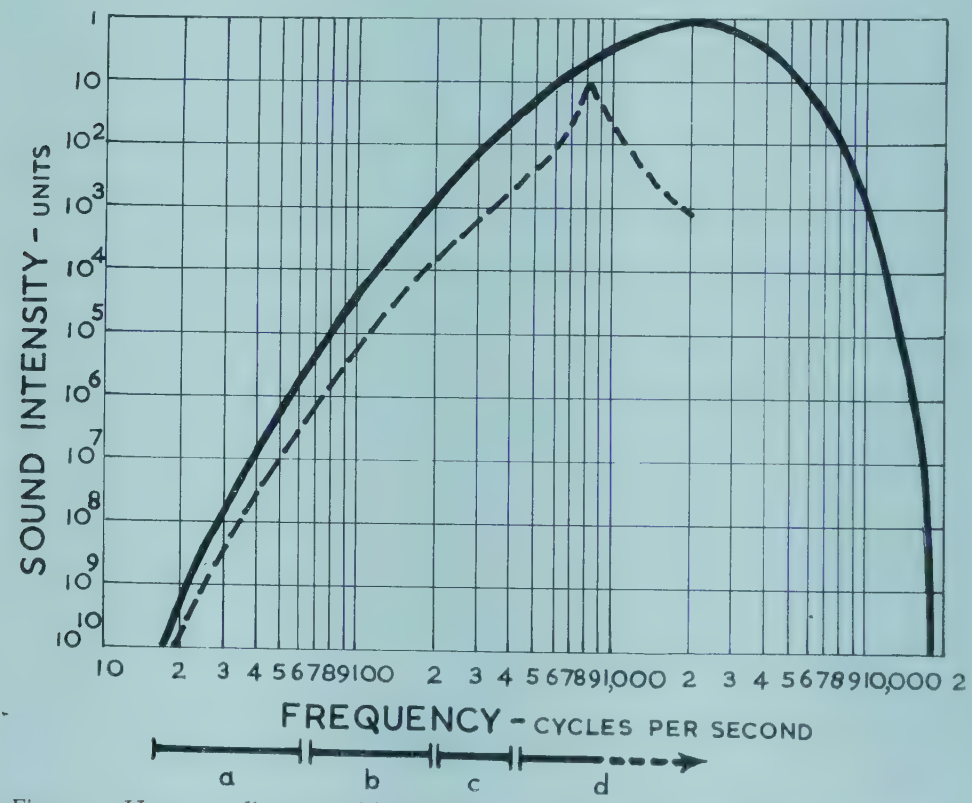


Figure 2. Human audiogram with superimposition of the frequency response curve of the 'high frequency' channel of the phonocardiograph (dotted line) as originally installed at the London Hospital and the Heart Hospital. The peak at 800 cycles and the fall off in sensitivity above this frequency were later 'improved', but did not in practice prove deleterious. The response of the microphone was tested separately and shown to be almost flat between 30 and 800 cycles.

cutting may occasionally be valuable; it is also an advantage to have one or two other filters which produce less severe attenuation of low frequencies and are therefore useful for emphasizing low frequency sounds and murmurs (Fig. 3). The galvanometers should have a frequency response which is flat to over 1,000 cycles and the recording must be made photographically on film or paper in order to achieve the necessary high frequency response and sufficient sharpness of detail at convenient paper speeds.

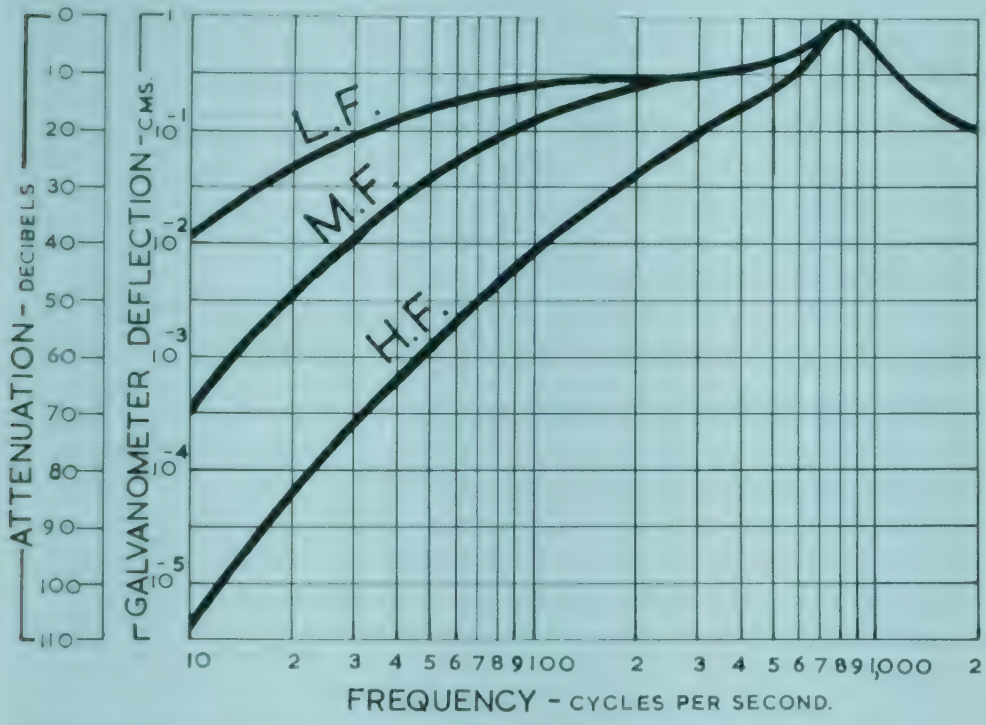


Figure 3. Frequency response curves of our low (LF), medium (MF) and high frequency (HF) channels.

A paper speed of 100 mm. sec. is sufficient for most purposes and avoids great lengths of recordings which are expensive and inconvenient to analyse. A quiet room is essential, but complete sound-proofing is not necessary at the moment since the limiting factor in high gain recordings is usually noise inherent in the patient.

The biggest single factor in achieving satisfactory recordings lies in careful attention to details by the operator. A suitable patient must be chosen; unrelaxed muscles produce a noisy base-line (Fig. 4). It follows that the patient must be at ease on a comfortable couch. The microphones must be exactly sited at the point of maximum intensity of the wanted

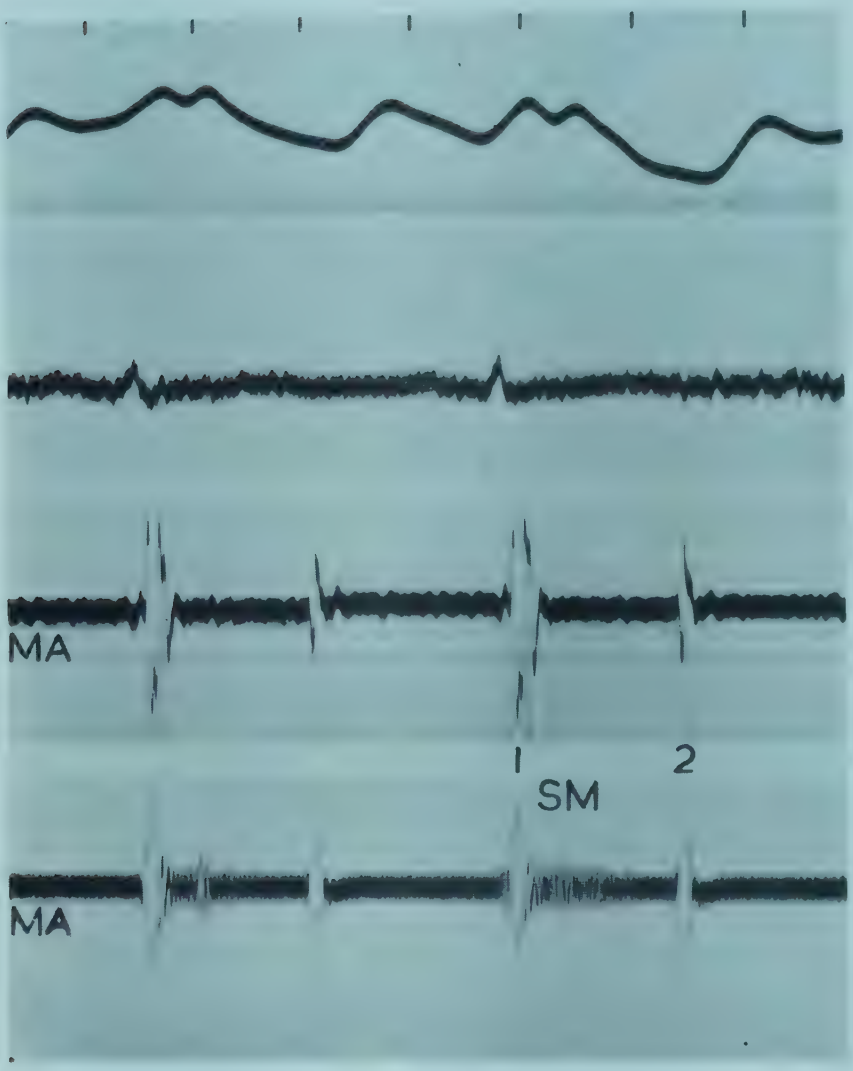


Figure 4. Bad base-line due to imperfect relaxation of patient.

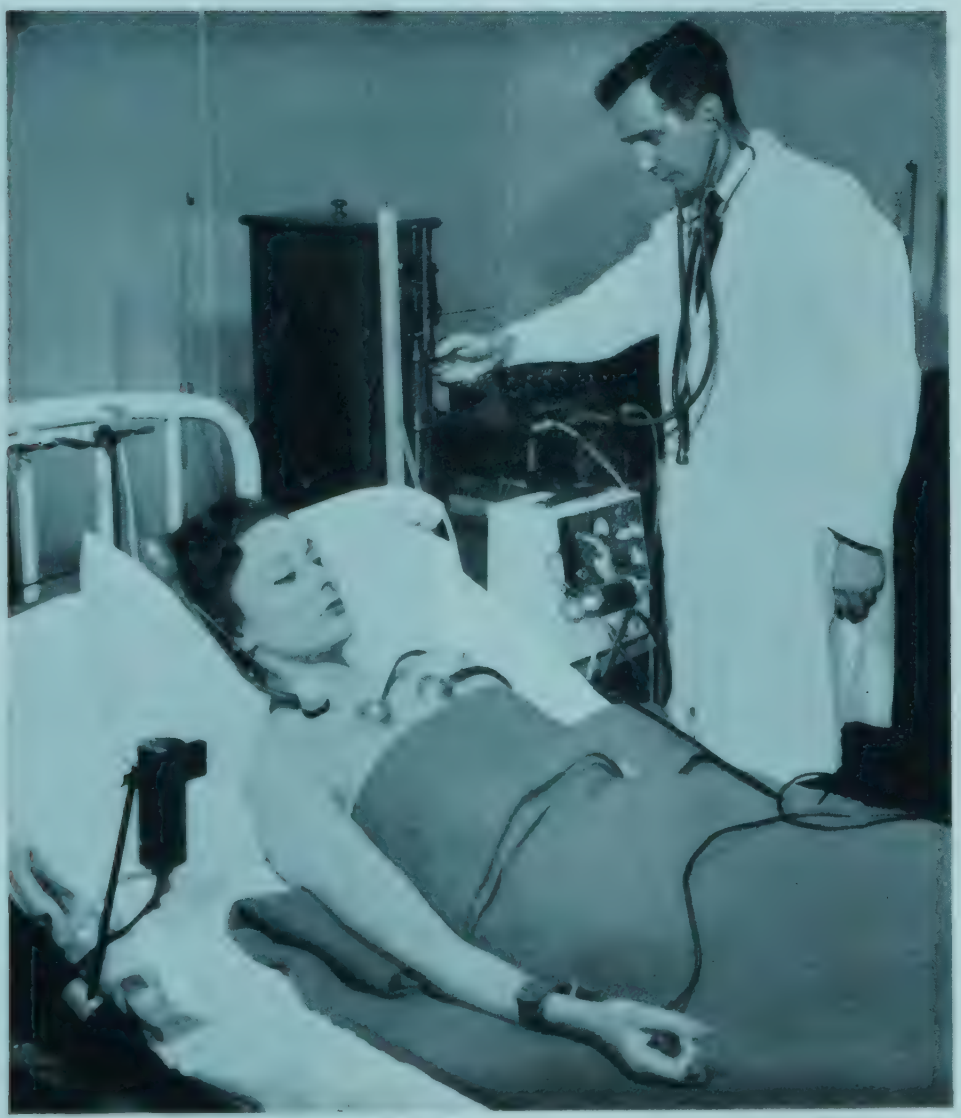


Figure 3. The patient relaxed and microphones placed accurately at the points of maximum intensity of the sounds and murmurs to be recorded. Respiration is recorded with a photocell which is stimulated by a beam of light which is interrupted by inspiratory movements of the chest wall.

sound or murmur (Fig. 5). In order to interpret the recording it is essential to have reference tracings. Another phonocardiogram from a different area together with an indirect carotid tracing and an electrocardiogram have been found to be the best combination, but necessitate a four-channel recorder (Fig. 6).

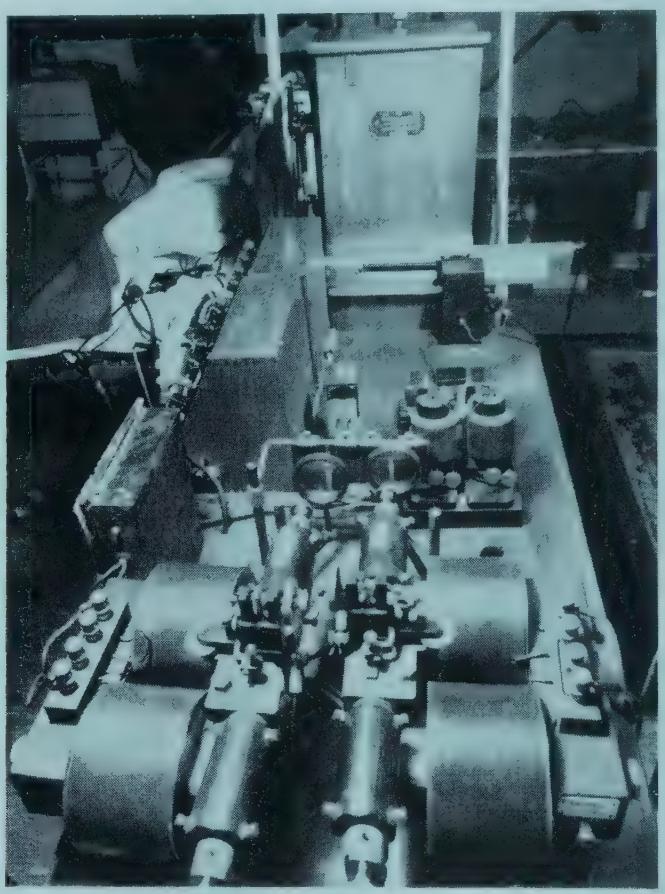


Figure 6. Apparatus installed at the Heart Hospital. In the foreground on the left are two string galvanometers, on the right two mirror galvanometers, and the time marker and camera are at the far end of the table.

Application of Phonocardiography in Medicine

(1) Research into the relation of sounds and murmurs to each other and to the pressure pulses.

(2) Routine diagnostic problems.

(3) Teaching. Learning to visualize the phonocardiogram would seem the best basis for teaching auscultation to undergraduates. In practice the

phonocardiogram is simply drawn on the blackboard. For training cardiac physicians there is no better way than by actually carrying out phonocardiography combined with clinical auscultation and evaluation of all the other cardiac findings.

(c) Research. There is only time to discuss the more recent applications

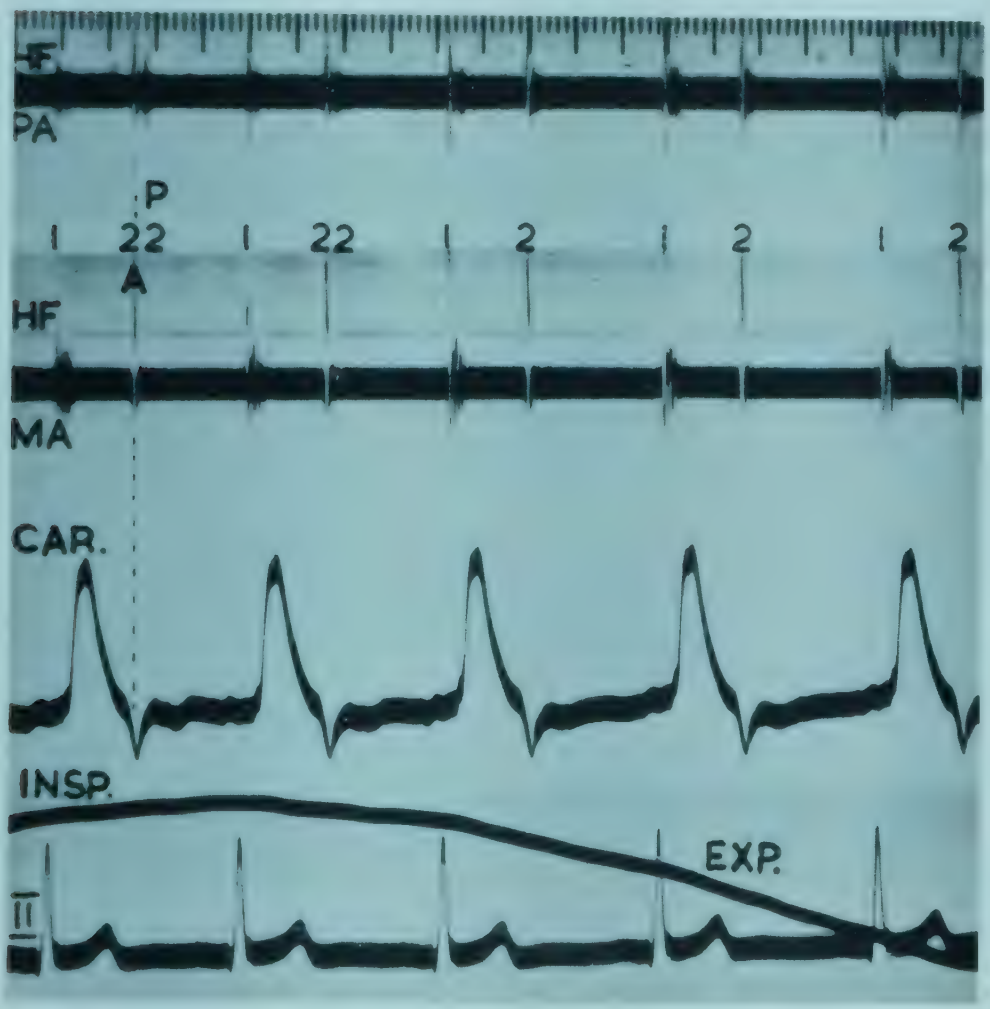


Figure 7. Physiological splitting of the second heart sound confined to the inspiratory phase of continued respiration. Aortic closure precedes pulmonary, and only aortic closure is normally transmitted to the mitral area.

of phonocardiography to clinical research. One of the foremost is perhaps a better understanding of the second heart sound. Potain was well aware that the second heart sound in the pulmonary area split into two components during the inspiratory phase of respiration but he could not obtain proof of their identity and even these facts seem to have been lost

sight of. Interest in the second heart sound was revived in London soon after the last war when splitting of the second sound was noticed by clinical auscultation. Phonocardiography revealed that in normal subjects the first component was due to closure of the aortic valve and the second to closure of the pulmonary valve (Fig. 7). The splitting was confined to the pulmonary area, or nearby, since the pulmonary component was too soft to be transmitted far beyond that area. The second heart sound at

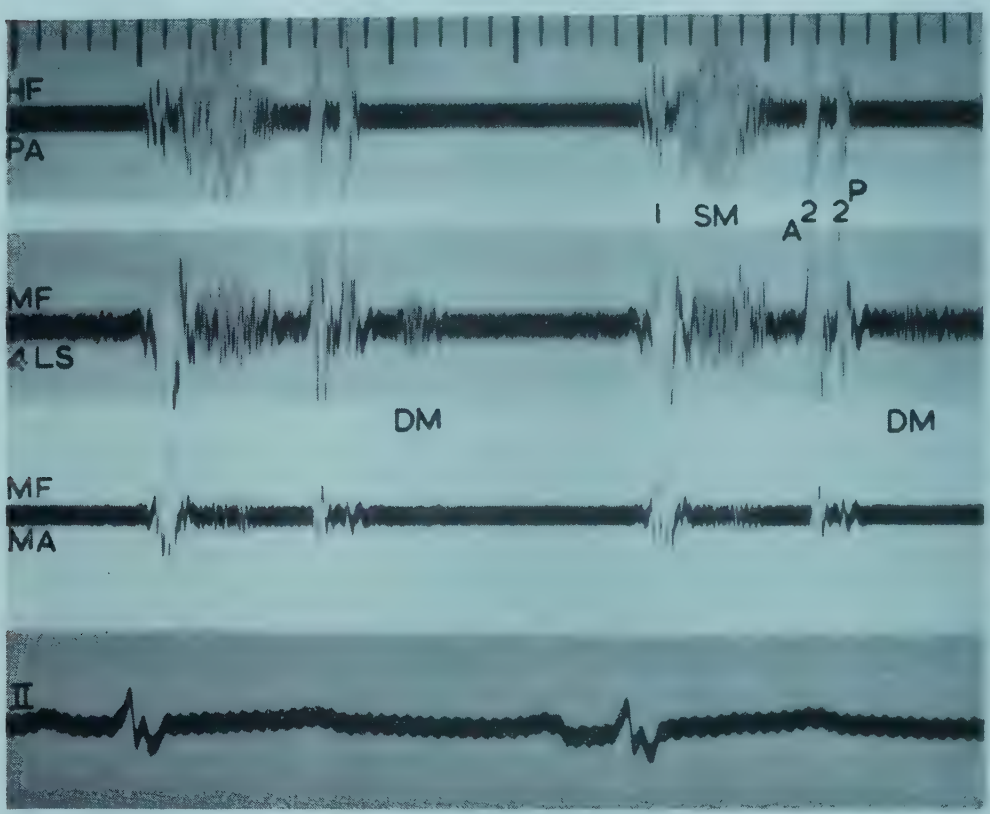


Figure 8. Atrial septal defect (with left to right shunt). Splitting of the second sound varies little with respiration (fixed) remaining wide in the expiratory phase of respiration and thus easily distinguished from the normal. The diastolic murmur is seen to start appreciably after pulmonary closure and is ascribed to rapid right ventricular filling consequent upon the increased right sided flow, the latter also being responsible for the short pulmonary ejection systolic murmur. The time intervals in this and subsequent records are 0·04 and 0·2 secs.

the mitral area was found to be solely due to closure of the aortic valve. It is most important to remember that aortic and pulmonary closure are almost simultaneous during the expiratory phase of *continued* respiration and only split during the inspiratory phase. In the latter phase the splitting may however be wide in normal subjects (up to 0·1 sec.). Resulting from

this knowledge of the behaviour of the second heart sound in normal subjects, it was soon found that abnormalities of splitting of the second heart sound were invaluable in clinical diagnosis, particularly the wide fixed splitting of the second heart sound in atrial septal defect (Fig. 8), the delay in pulmonary closure in pulmonary stenosis, and reversed splitting of the second heart sound when there was delay in closure of the aortic valve as in left bundle branch block.

Once the behaviour of the aortic and pulmonary components of the second sound had been appreciated it became easy to study the opening snap of the mitral valve in mitral stenosis, for previously this sound had

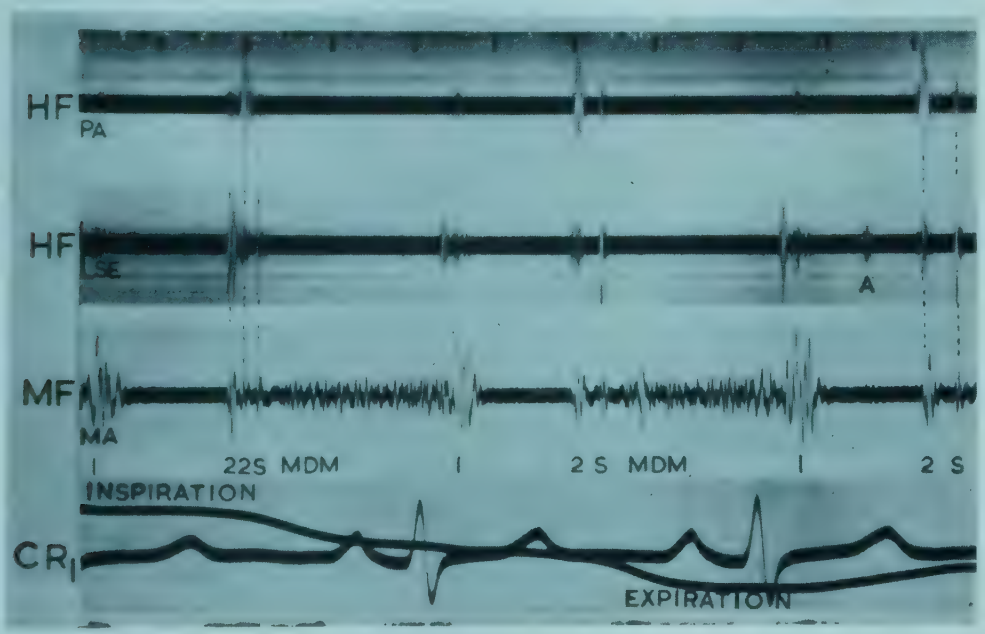


Figure 9. Mitral stenosis. Simultaneous phonocardiograms from the pulmonary area, lower left sternal edge, and mitral area. In the inspiratory phase of respiration aortic closure, pulmonary closure, and the snap are recorded.

been confused with that due to pulmonary closure. With simultaneous phonocardiograms from the pulmonary area and lower left sternal edge, both aortic closure, pulmonary closure, and snap were easily recorded (Fig. 9) and the very high incidence of a snap in mitral stenosis was then appreciated.

Attention was also directed to the first heart sound. Again Potain had noticed that in normal subjects two major components of the first sound could be identified by simple auscultation at the lower end of the sternum. By relating the sounds to central aortic tracings we found that the first

component was almost certainly due to closure of the mitral valve and the second to closure of the tricuspid valve. In many patients with heart disease, however, it was noticed that there was in addition another sound in early systole occurring *after* closure of the mitral and tricuspid valves. This had been noticed by Lian in France and had been described by Wolferth as the 'semi-lunar opening click'. It was found that this sound occurred in association with dilatation of the aorta or the pulmonary artery. From indirect carotid tracings, and pressure pulses from the pulmonary artery, it was shown that this sound occurred shortly after ejection of blood into the great vessels and was therefore called an ejection sound. The pulmonary variety was usually associated with pulmonary hypertension and the measurements of its timing were taken under these circumstances. A similar sound also occurred with both aortic and pulmonary valve stenosis and was presumably also related to ejection into the area of post-stenotic dilatation, but the exact timing of this sound in relation to the pressure pulse from the great vessel was impossible owing to distortion of this pulse by the stenosis. It has proved to be a valuable physical sign in differentiating valvar and sub-valvar pulmonary stenosis, and may prove to be of equal value in relation to the aortic valve. Other ways in which phonocardiography has been of value in deciding the timing of sounds and murmurs includes investigation of the diastolic murmur so frequently encountered in atrial septal defect (Fig. 8). Once the mistake of ascribing the murmur to additional mitral stenosis had been appreciated, the murmur was then thought to be due to pulmonary regurgitation. Pulmonary regurgitation, however, is extremely rare without pulmonary hypertension, and most patients with atrial septal defect have of course low pulmonary pressures. With phonocardiography it was soon found that there was a gap between pulmonary closure and the onset of the murmur (Fig. 8) which is now ascribed to turbulence produced by the rapid flow of blood through the tricuspid valve; but in atrial septal defect complicated by high pulmonary vascular resistance, a true pulmonary diastolic murmur occurs; it is associated with a single second heart sound and the tricuspid diastolic murmur is absent, for the left to right shunt has diminished.

Phonocardiography has also been invaluable in differentiating ejection systolic murmurs from regurgitant murmurs. Ejection of blood into the great vessels when associated with stenosis, increased stroke volume, valvular deformity or a combination of these factors, produces a short systolic murmur which rises to a crescendo about mid-systole and finishes before the second heart sound (Fig. 8). Regurgitant murmurs such as

those due to mitral or tricuspid regurgitation are always pansystolic. Attention should be specially directed to the end of systole where a regurgitant murmur is still present when it would have ceased if the murmur had been due to ejection into the great vessels (Fig. 8). The explanation of the continuation of back-flow and of murmur in late systole is simply that there is still a large pressure difference between the ventricle and atrium. This has proved to be a valuable distinction since

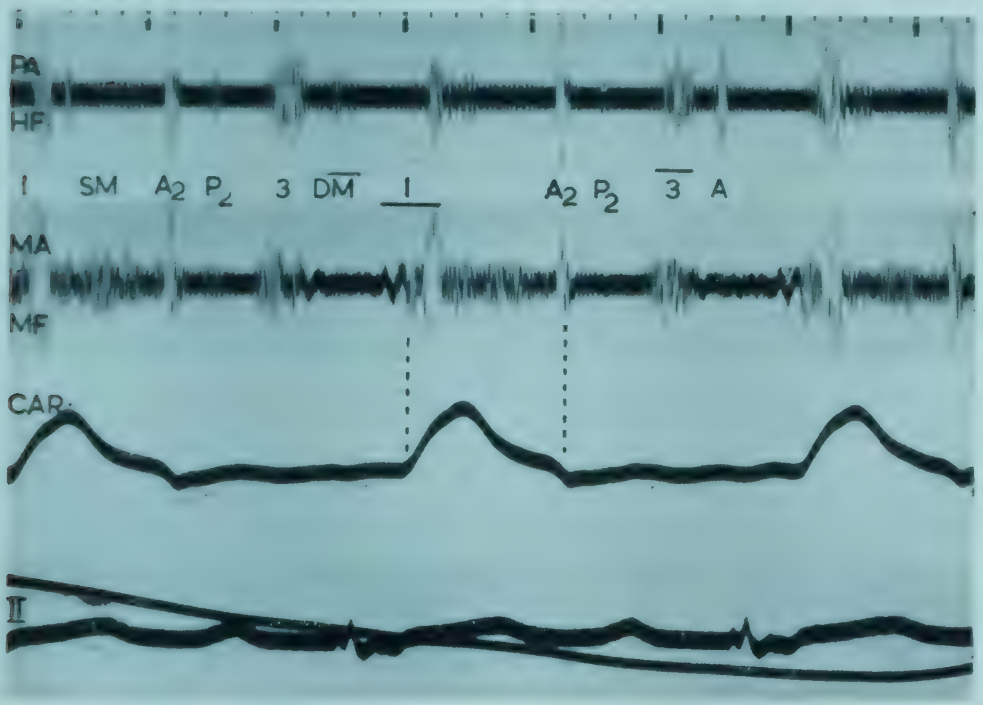


Figure 10. Ebstein's anomaly. The first sound is split and the loud later component is probably due to tricuspid closure. The pansystolic murmur is ascribed to tricuspid regurgitation, the wide splitting of the second sound to right bundle branch block, and the third heart sound and diastolic murmur to abnormalities of right ventricular filling.

aortic murmurs are frequently transmitted to the mitral area making differentiation difficult from mitral murmurs.

It will be seen from these examples that phonocardiography has been of value mainly in deciding the timing of sounds and murmurs and relating them to each other and to the pressure pulses. Once the relation of sounds to murmurs has been demonstrated by phonocardiography, most points can be picked up by simple clinical auscultation, though in complicated cases a phonocardiogram may still be useful (Fig. 10).

Phonocardiography has been of little value for picking up faint murmurs, for the ear is an extremely sensitive organ and difficult to beat with phonocardiography. Furthermore our hearing mechanism is able to concentrate and discern pitch and quality while the phonocardiogram records all vibrations which are present.

ACKNOWLEDGMENTS

I am grateful to the editor of the *Postgraduate Medical Journal* for permission to reproduce Figs. 1 and 2, the editor of *John Bull* for Fig. 5, the editor of the *British Heart Journal* for Fig. 8, the editor of the *Lancet* for Figs. 7 and 9, and, above all, to many friends and helpers without whom this work would not have been possible.

REFERENCES

LEATHAM, A. (1958) *Lancet*, ii, 703 and 757.
LEATHAM, A. (1958) *Paed. Clin. N. America*, November.

VECTORCARDIOGRAPHY

Wallace Brigden

Cardiac electrophysiology has become a large subject since Waller made the first recordings of the human electrocardiogram in 1887. Contributions to this field of biological knowledge have come from electrical engineers, physicists, biophysicists, physiologists and clinicians. Observations have been made on isolated cells, muscle preparations, whole heart perfusions and intact experimental animals. Empirical information obtained from the analysis of electrocardiograms from patients has been steadily collected, and its value enhanced by correlation with necropsy examinations. In the time available it is not possible to evaluate the research in this large field. It is, however, particularly relevant to indicate that many of the major advances have been due to technical developments in the tools of research and modern medicine is especially indebted to physicists and technicians in this field. The construction of the string galvanometer and its subsequent modification by Einthoven enabled him to pioneer modern clinical electrocardiography. A great expansion of investigation followed the introduction of valve amplifiers and led to the recording from local areas on the chest wall. It is known that Einthoven and his associates (1913) were thinking in terms of vectorcardiography, but it was the development of the cathode ray oscilloscope which facilitated the recording techniques introduced independently by Schellong in 1936 and Wilson & Johnston in 1938.

The metabolism of cardiac cells results in such properties as rhythmicity, contractility, excitability and refractoriness, and the production of an electrical current is inseparable from these vital functions. We are not here concerned with the fundamental behaviour of heart muscle cells, but rather with the projection of the total electrical effects of this muscle through the body tissues to the surface and its graphic recording there. It is necessary, however, to mention some basic concepts concerning the source of the electric current and of the medium through which it is transmitted, for it is upon these concepts that the basis of vectorcardiography rests.

The Dipole Theory

Cell activity in a conducting medium results in a potential difference being built up between boundaries of active and relatively inactive tissue. The transfer of electrons results in current flow. Positivity resulting from activity is closely related to areas of negativity; there is no surplus, no generation of electricity but only the transient build-up of potentials in a medium which permits current flow and subsequent rapid rebalance. All of the point sources of activity in a single fibre summate. Further aggregation of the effects of each and all of the fibres results in a total

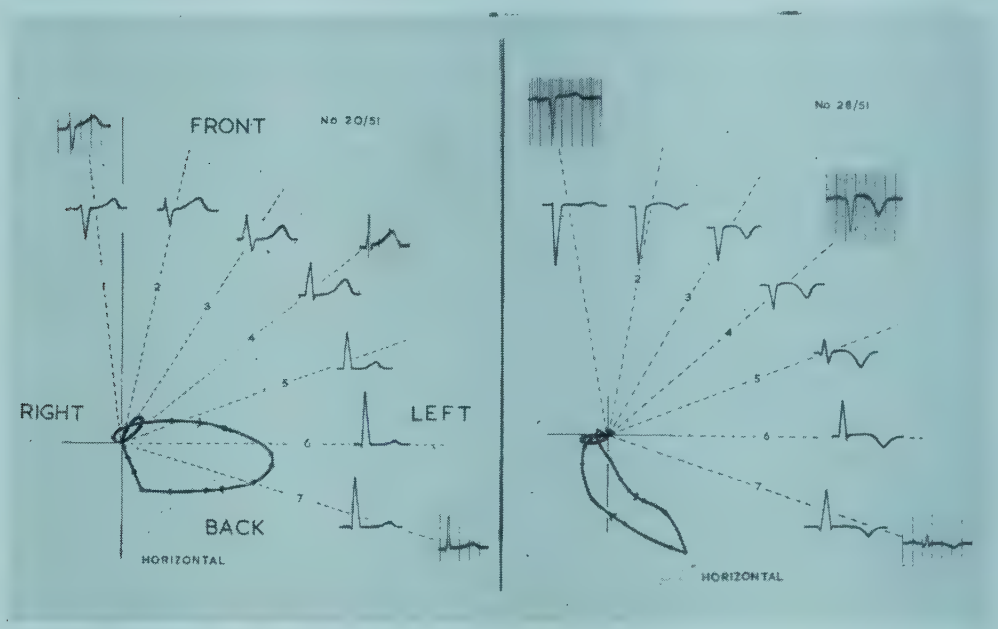


Figure 1. The horizontal projection of the vectorcardiogram of a normal adult on the left and of a patient with anterior cardiac infarction on the right. Reconstructions of the scalar electrocardiogram on each of seven axes (in the positions V.1 to V.7) are shown together with the electrocardiogram recorded at V.1, V.4, and V.7 positions.

electrical effect for the whole heart which may be assumed to act as a single dipole. Such a concept is probably no more than an approximation which is convenient in the study of human physiology, but it is an approximation which fits many observed phenomena when the heart is relatively small. The central source may be expressed at any one instant as a composite vector, having magnitude, direction and sign. Such a composite heart vector is a manifestation of all of the activity passing in different directions through the myocardium and the heart cavities at

that time (Robertson, 1951) and it represents no more than the dominant trend of these multiple electrical activities. It is clear that the electrocardiogram or vectorcardiogram can only be accepted as recording the resultant of all dipoles if the heart behaves electrically as a central point source or single dipole. Viewed anatomically this may seem unreasonable when one considers the large size of the organ, its different parts, its irregularities and its eccentric position in the conducting medium; yet there is much to support the view that the heart behaves approximately

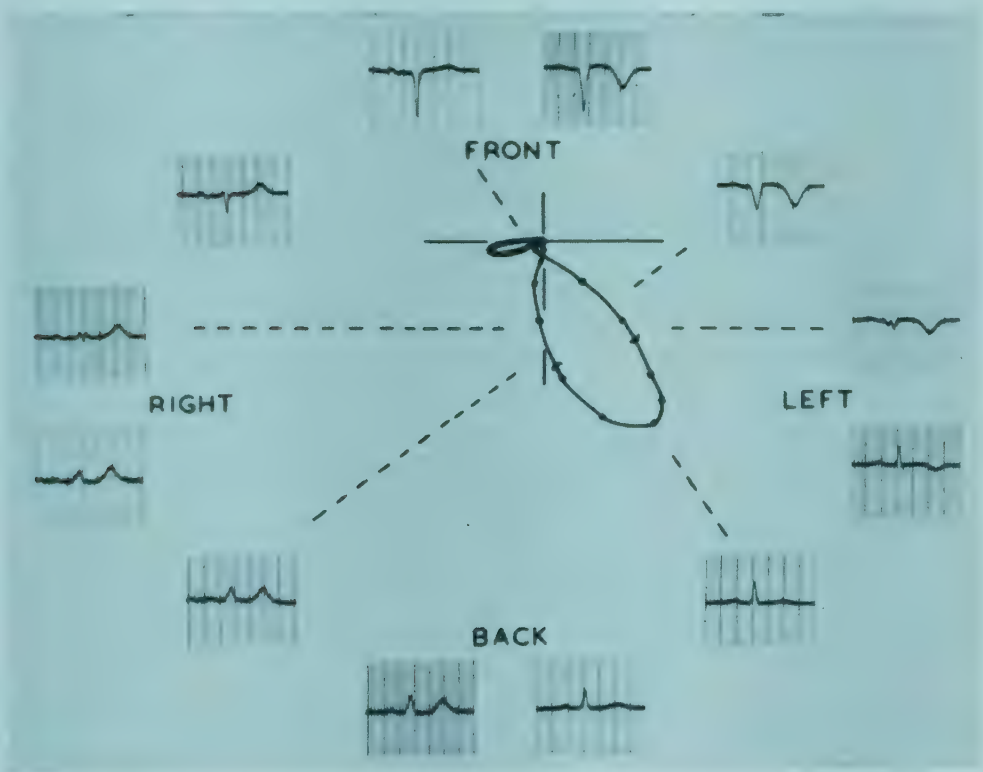


Figure 2. Horizontal projection of the vectorcardiogram in anterior cardiac infarction shows the loss of anterior vectors causing a dorsal position of the loop, and Q waves in V.1 to V.6. Electrocardiograms recorded at opposite ends of axes through the chest show 'mirror patterns'.

as a single dipole which can be represented as a single resultant vector at any moment.

The validity of the single equivalent dipole depends on evidence from appropriate models of the thorax and from the following observations in man.

(*a*) Vectorcardiograms obtained from different lead positions (reference

systems) show a high degree of agreement when certain lead coefficients are introduced.

(*b*) Similarly, electrocardiograms reconstructed from vectorcardiograms are qualitatively similar to actual tracings recorded in the conventional way (Fig. 1).

(*c*) Scalar electrocardiograms recorded at opposite ends of axes through the heart show remarkable mirror patterns (Schmitt, Levine & Simonson, 1953) (Fig. 2) and when allowance is made for the effect of distance upon the more remote electrode, cancellation may be complete. As might be expected, a failure of cancellation is more likely to occur when the heart is abnormally large.

The Conducting Medium

It is generally assumed that the body tissues between the heart and the surface behave as a homogeneous medium. However, on *a priori* grounds it would seem unlikely that the various tissues in the thorax have the same conducting properties and any such difference might be expected to introduce error in attempts to express the central electric source in terms of surface leads. Indeed Burger & Van Milaan (1946, 1947 and 1948) have shown that there are differences between lung and other tissues. However, in practice this inhomogeneity seems to be rather unimportant except perhaps for the low resistance and consequent short circuiting effect of blood within the heart cavities. The surface potential at a distance from the central source varies inversely as the square of the distance, and depends on the angle subtended by the site of the recording electrode to the axis of the cardiac vector. Thus differences between lead patterns depends on the distance of the electrode from the source, the angle it subtends with the axis of the vector and the nature of the underlying conducting tissues.

The short circuiting effect of blood within the heart presumably diminishes the effect of all the separate sources and leads to a unification of electrical activity, which lends further support to the hypothesis that the heart behaves as a single dipole.

At any instant the level of dipole activity is expressed by a resultant vector having magnitude direction and sign. For the whole cardiac cycle there will be an infinite number of vectors whose extremities, representing magnitude, when conjoined will form a continuum which is the spatial vectorcardiogram. This is a figure having three dimensions which may be reconstructed or visualized from its plane projections. Each projection

(horizontal, frontal and sagittal) may be obtained by the integration of a pair of appropriate co-ordinates—a pair of bipolar or unipolar leads representing vertical, transverse, or antero-posterior potential differences (Figs. 3*a* and 3*b*). The cathode ray oscilloscope is convenient for recording these potential differences. The minute electrical forces involved require amplification by 100,000 to 200,000 times before sufficient beam deflection is produced. The potential difference applied to the vertical plates deflects the beam laterally and that applied to the lateral plates deflects it vertically; when both sets of electrodes are used simultaneously the tube acts as an integrator by shifting the beam to a spot which is proportional to the geometric sum of the two co-ordinates. The beam may be interrupted at an appropriate frequency to provide an accurate time-marker; its direction is indicated by imposing a tear-drop configuration to the beam by a suitable oscillator.

Scalar electrocardiograms recorded from two pairs of electrodes may also be integrated by means of a simple mechanical drawing instrument (Shillingford & Brigden, 1952). For those experienced in the art of inspection and visual integration of two scalar electrocardiograms no vectorcardiogram is necessary!

Reference Systems

A reference system is a three-dimensional figure which can be defined by a series of points (electrode positions) on the surface of the body. Since the absolute cardiac vector is unknown, the conducting medium inhomogeneous, and the heart rather large in relation to the body, no system provides more than a rough approximation to the cardiac vector in man. The inadequacy of the various systems is reflected in their multiplicity.

The selection of types of lead, lead positions and polarity has caused much discussion. Even if investigators differ about lead placing, at least polarity should be consistent. The same convention of sign should be used for vectorcardiography as electrocardiography as there is no fundamental difference between the two methods (Robertson, 1951).

Representation of the spatial vector loop requires a record of its projection on to two or more planes. A plane projection is derived from two common components of the reference system. The leads may be unipolar, bipolar or a combination of both.

Einthoven's triangle has been widely used as the basis of an *equilateral tetrahedron reference system*. The triangle represents the frontal plane and a unipolar electrode placed on the back provides a record of depth so that

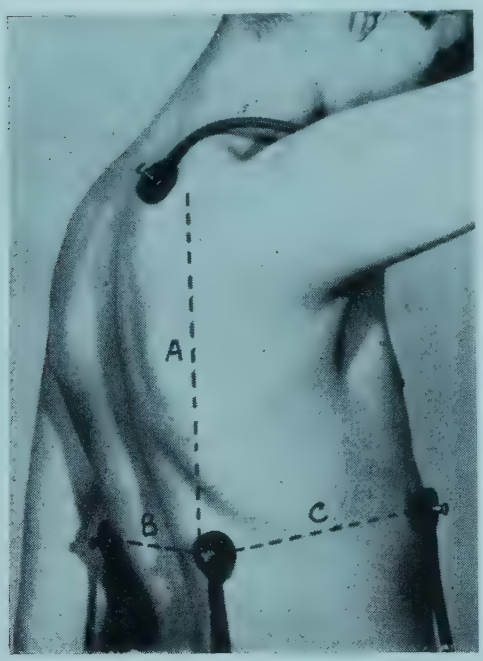


Figure 3a. Bipolar lead positions for recording the vertical (A), transverse (B) and sagittal (C) components of an orthogonal reference system.

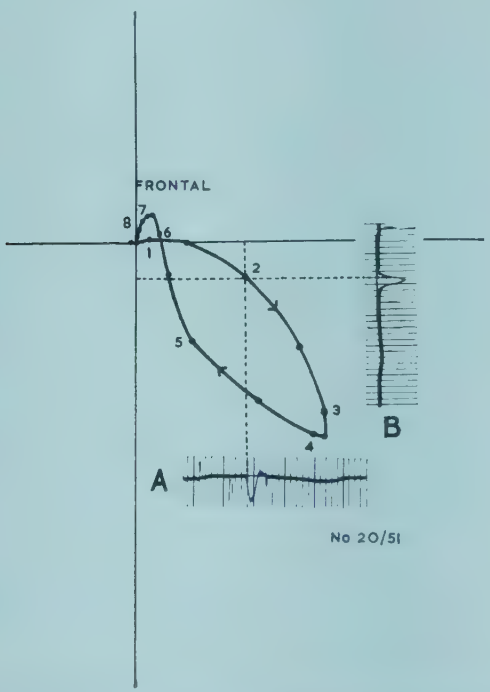


Figure 3b. Vertical scalar electrocardiogram *A* integrated with transverse lead *B* by means of a drawing instrument. This forms the frontal projection of the vectorcardiogram.

sagittal and horizontal components may be obtained. Wilson & Johnston derived the frontal plane from lead I which is arranged to give a horizontal deflection to the oscilloscope spot, and VF to produce a vertical deflection. The posterior point of the tetrahedron system is placed at the level of the seventh thoracic spine 2 cm. to the left of the mid-line. It has been shown that this point is roughly as remote as the extremity leads from a dipole placed near the heart (Abildskov, Burch & Cronwich, 1950).

Burgher & Van Milaan have shown that the triangle is not electrically equilateral but they considered that it was more accurate than some rectangular systems. Grishman and others (1951 and 1952) point out that distortion from the triangle system is increased when it becomes the basis of a spatial system. It is clear that extremity leads give a tilted record of the frontal plane, and furthermore the integration of lead I and a unipolar lead on the back provides an even more oblique 'view' of the horizontal plane. Whilst there is no entirely satisfactory reference system many workers have now adopted some type of orthogonal system (Fig. 3*a*).

Schellong (1936) introduced an orthogonal system using three bipolar leads on the thorax, placing the common electrode at the left shoulder. This position has been shown to be unsatisfactory for vectorcardiography; and Schaffer (1957) showed that the left arm lead (same as the shoulder) is especially sensitive to changes in electrical position of the heart. Duchosal & Sulzer (1949) introduced a different orthogonal system using three bipolar leads with an electrode on the right lower back producing a long vertical component. This double cube or parallelepiped system causes an unequal representation of electrical forces. Furthermore, these authors have used negative polarity whereas most workers have accepted Einthoven's positive convention of sign, that is in a 'minus-plus' direction so that the origin of the vector is negative and the 'wandering' end position positive. Consistent convention of sign has the further advantage that vectorcardiograms are more easily related to the conventionally recorded electrocardiograms. The single cube system of Grishman, Borun & Jaffe (1951) has some advantages and is being widely used. The electrodes are placed at the corners of a cube: the common electrode is placed in the right posterior axillary line at the level of the first and second lumbar vertebrae. Positivity is recorded downwards, forwards and to the left.

In a system described by Donzelot, Milovanovitch & Kauffman (1950) V.2 and V.6 are used for sagittal and transverse components, but the close proximity of these electrodes to the electrical centre of the heart introduces distortion. The precordial electrodes are so close that the anterior myocardium exerts a great influence on these leads compared with other parts

of the heart. Furthermore, V.2 and V.6 are not in the same horizontal plane. These authors like Duchosal & Sulzer used negative polarity. The method is not widely accepted.

The multiplicity of reference systems is often used to indicate the essential imperfections of the method. However, in a broad way the consistency of findings derived from the different routes is more impressive than the minor variations, and such consistency indicates a fundamental validity of the vector approach to cardiac electrophysiology.

The Normal Vectorcardiogram and its Relation to the Electrocardiogram

At any instant the result of all cardiac electrical forces may be represented as a vector orientated in space, having magnitude, direction and sign. Instant vector follows instant vector presenting a continuum: acceptance of the dipole hypothesis indicates a central constant point of origin and an infinity of distal points which when joined form a loop. The spatial vectorcardiogram during the cardiac cycle represents atrial depolarization and repolarization followed by ventricular depolarization and repolarization. In practice it is technically difficult to record atrial activity unobscured by following ventricular activity. At best a small loop of atrial depolarization (the *P* loop) can be recognized; this is followed by a much larger loop of ventricular depolarization ($QRS\hat{E}$) and a small loop of ventricular repolarization (the $T\hat{E}$ loop). Each loop is closed and, during the brief intervals between the cessation of one and the onset of the next, there is no significant changing electrical activity. The vectors have no magnitude, the oscilloscope spot is at zero and is equivalent to the isoelectric *P–R*, *S–T* and *T–P* intervals of the electrocardiogram.

In the ventricular phase it is generally considered that the left side of the septum is activated first, and this produces vectors of small magnitude directed anteriorly and to the right (see Fig. 1). Scalar leads on the left side of the chest record this as the initial negative normal Q. Activation of the main mass of the left ventricular muscle produces overwhelming activity downwards and to the left resulting in a bold vector loop in this direction: which is equivalent to large *R* waves and deep *S* waves in left and right chest leads. The first and last parts of the loop are inscribed more slowly than the remainder. Using positive polarity and the reference system described, the normal loop is clockwise in its frontal plane projection, counter-clockwise in the horizontal plane and clockwise in the sagittal plane.

The *T* loop is smaller than the *QRS* loop but it is orientated in the same general direction. The relationship is similar to that between the *QRS* and *T* complexes of scalar electrocardiography. The relationship between the appropriate electrocardiographic leads and the plane projections of the spatial vectorcardiogram is shown in Figs. 1 and 2. A unipolar lead may be regarded as a recording of the total forces at any one point projected on to the axis of a line joining the dipole centre to the surface electrode. The unitary nature of electrocardiogram and vectorcardiography is illustrated by the remarkable degree of consistency in the lead patterns reconstructed from a vectorcardiogram obtained from remote and entirely different electrode positions and the actual scalar leads recorded from the conventional surface positions.

The Abnormal Vectorcardiogram

The electrocardiogram provides two types of information. Firstly, as an exact time-marker, the relations between one heart cycle and the next are recorded; thus arrhythmias may be clearly recognized. This kind of information, however, may be obtained from clinical examination and other recording methods. The second type of information from cardiac electrophysiology concerns the sequence and level of electrical activity throughout the myocardium during one beat; this kind of information has no direct clinical counterpart. A pathological Q wave cannot be palpated, it is not audible, nor can it be recorded from the pulses. In vectorcardiography we are mainly concerned with the second type of information but arrhythmias may be studied by a ciné film-recording of the sequence of loops on the oscilloscope.

Vectorcardiographic studies have now been made in many kinds of heart disease; however, much further research and experience is required before this method can be fairly evaluated and compared with scalar cardiography.

CARDIAC INFARCTION

The main vectorcardiographic abnormalities in cardiac infarction have been reported extensively (Duchosal & Sulzer, 1949; Scherlis & Grishman, 1951; Brigden & Shillingford, 1952; Grant & Murray, 1954; Burch *et al.*, 1955 and 1956; Milnor, 1958; Burch *et al.*, 1958). The vectorcardiogram and electrocardiogram in cardiac infarction are products of the activity of surviving cardiac muscle since necrotic tissue is inert. The cardiac vectors, which are the instant resultants of all electrical forces, are deviated into

abnormal positions away from the site of the infarct because of the unbalanced activity of good muscle.

Any part of the spatial loop may be abnormal, but the early vectors (0·01–0·03 sec.) are most commonly affected. The mid-vectors are also frequently displaced giving the loop an indented appearance. Abnormalities in the terminal portion of the loop are rather uncommon and less obvious than in other parts of the loop. The direction of rotation is sometimes abnormal in one or more of the plane projections.

The vectorcardiogram of *anterior infarction* is determined by the loss of early and anteriorly directed vectors. Thus the loop lies in a more posterior position than normal and also deviates upwards and to the left. These abnormalities are best seen in the sagittal and horizontal planes (Figs. 1, 2 and 4) and in the precordial leads of the electrocardiogram. In *diaphragmatic or posterior infarction* the loop is deviated in a cranial direction due to the influence of unopposed vectors arising in undamaged anterior and superior parts of ventricular wall. Infarction in this site is thus best seen in the frontal and sagittal plane projections of the vectorcardiogram and in standard leads, 2, 3 and unipolar lead VF of conventional cardiography (Fig. 5).

Vectorcardiography provides a more rational basis for the understanding and interpretation of scalar leads in cardiac infarction but there is no evidence that the vectorcardiogram is significantly superior to the electrocardiogram in this diagnostic field. Indeed, minor and local changes in *T* waves and *S–T* segments are not as readily recognized in the vectorcardiogram as in the electrocardiogram. It is probable that the short distance between a precordial exploring electrode and the site of injury exaggerates local effect so that this type of abnormality is recorded selectively.

VENTRICULAR HYPERTROPHY AND BUNDLE BRANCH BLOCK

The vectorcardiogram in myocardial hypertrophy is most obviously abnormal during depolarization. Distinctive features are found in the *QRS* loop which is abnormal in position or contour or both (Gardiner & Lowe, 1953 and 1954).

In *left ventricular hypertrophy* the vectorcardiogram shows a bold loop to the left, generally distorted backwards and more upwards than normal; these abnormalities are seen best in the horizontal and sagittal planes. The direction of rotation remains normal. However in *left bundle branch block* the direction of rotation is usually reversed in the horizontal loop and the

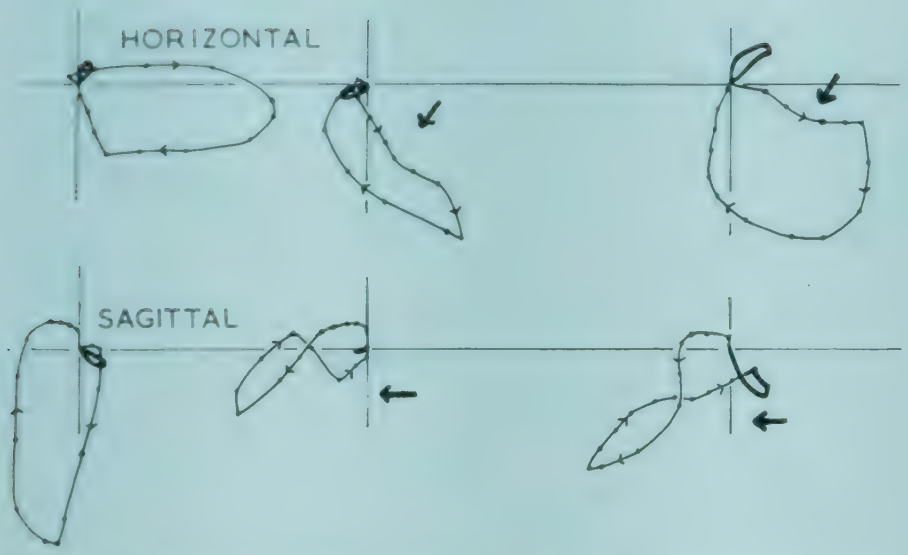
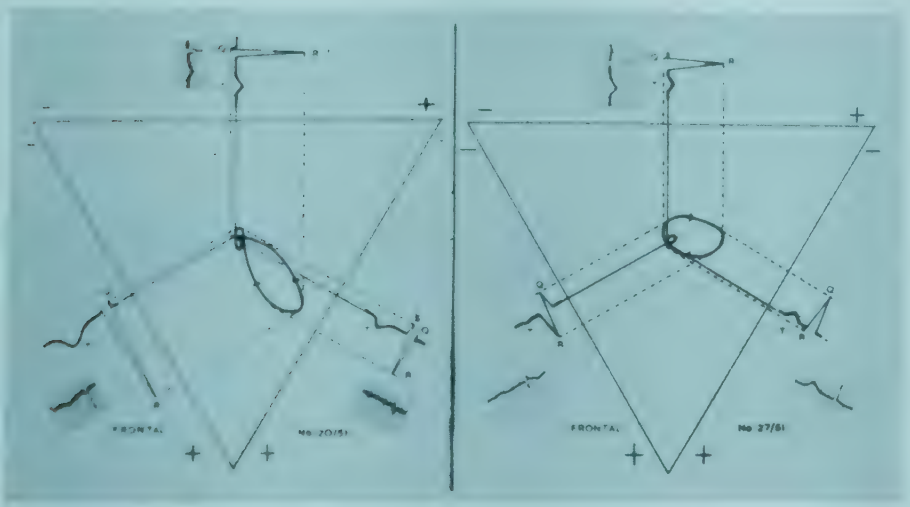


Figure 4. Horizontal and sagittal projections of the loop in two cases of anterior cardiac infarction show its dorsal position when compared with the normal loop shown on the left. (Arrows indicate 'dorsal drift'.)



(a) (b)

Figure 5. (a) The frontal projection of a normal vectorcardiogram with reconstruction of lead, I, II, and III, and with actual recordings for comparison. (b) The frontal plane vectorcardiogram of a case of posterior cardiac infarction. A deep QII and QIII are associated with the upward shift of this loop.

delayed conduction is shown by closely spaced time-markings. It is probable that the electrocardiographic diagnosis of left bundle branch block represents a range of conditions which at one extreme represents the electrical effects of the late stages of great hypertrophy, and at the other the effects of a small and possibly insignificant lesion of the septum without damage to the main muscle mass. It is probable that vector analysis may help to elucidate some aspects of this problem.

The vectorcardiogram in *right ventricular hypertrophy* and in *right bundle branch block* has been the subject of much investigation as these conditions

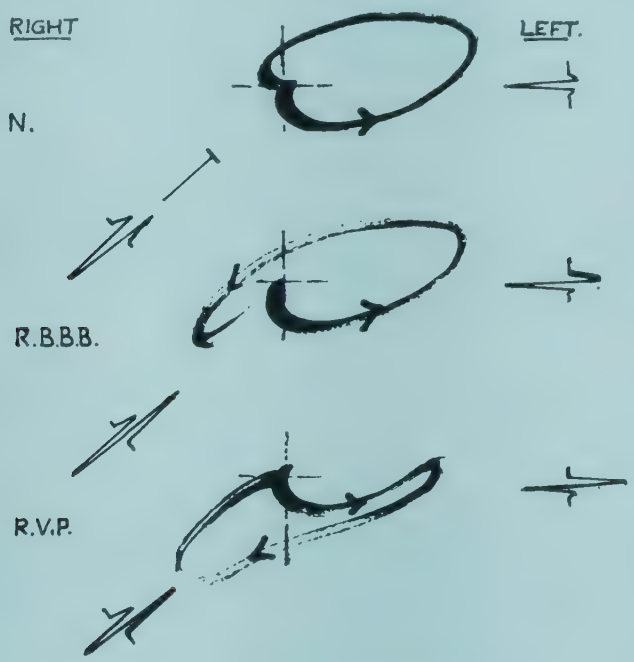


Figure 6. Diagrams of the horizontal projection of the vectorcardiogram in a normal adult (top), in right bundle branch block without hypertrophy, and in right ventricular hypertrophy. Superficially similar patterns may be produced at V.1 by right bundle branch block and right ventricular preponderance but the loops are different.

may present difficult diagnostic problems in conventional cardiography. Confusion lies in the fact that in isolated right bundle branch block without significant hypertrophy an *rSR* configuration is found in right pectoral leads, and a superficially similar pattern is often found in patients with right ventricular hypertrophy. The horizontal projection of the vectorcardiogram shows that there may be fundamental differences in qualitatively similar electrocardiograms; the major part of the loop may be normal with a brief terminal prolongation to the right whilst in right

ventricular hypertrophy the early vectors may be normally directed to the right and forwards followed by an overwhelming secondary deviation to the right (Fig. 6). Intermediate patterns (Fig. 7) and the influence of superimposed anatomical rotation may be more easily analysed by vectorcardiography; thus in mitral stenosis the difficulties in interpretation of right precordial leads lies in the combined influence of clockwise rotation of the heart about its vertical axis and hypertrophy of the right

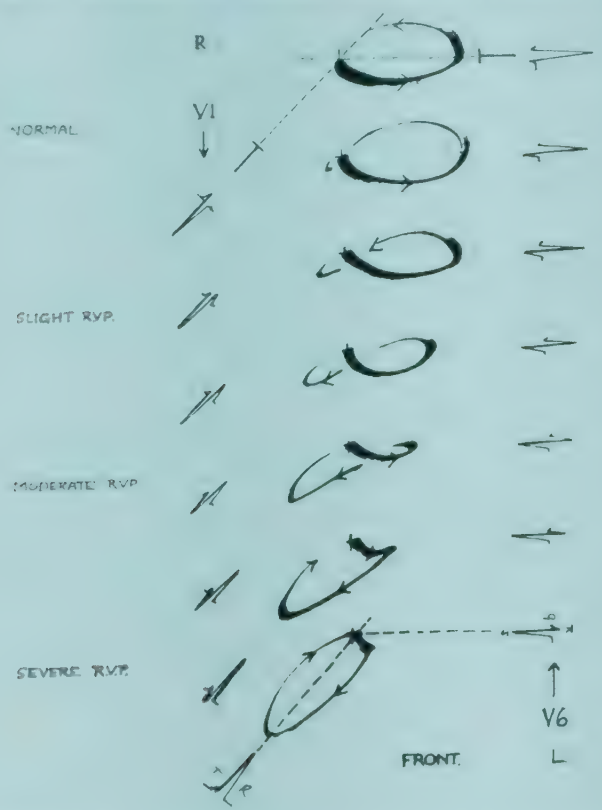


Figure 7. Diagrams of the horizontal projection of the loop showing the change from normal to an extreme degree of right ventricular preponderance with monophasic *R* waves in the V.1 position of the scalar cardiogram.

heart from pulmonary hypertension. The vectorcardiogram shows these different factors more clearly than the electrocardiogram (Shillingford & Brigden, 1954).

SOME CONCLUSIONS

1. The heart behaves as a single dipole and its activity throughout the cardiac cycle may be represented as a spatial vectorcardiogram. Scalar cardiograms are derivatives of this so that all the information provided

by any combination of electrocardiograms should be contained in the three plane projections of the vectorcardiogram, but in practice this is not so for at least two reasons. Firstly, precordial leads are relatively close to the myocardium, and show local abnormalities which are difficult to recognize in the vectorcardiogram. Secondly, when the heart is abnormally large it no longer behaves as a single equivalent dipole. However, the concept of the single equivalent dipole is useful for the majority of cases and its invalidity in large hearts implies that it is an oversimplification when applied to the whole range of heart disease.

2. The technique of obtaining vectorcardiograms is more difficult than the recording of an electrocardiogram, but analysis of the vectorcardiogram is no more difficult than interpretation of the electrocardiogram. The diagnostic implications of vectorcardiograms are less surely known at present and it is unlikely that vectorcardiography will replace the conventional electrocardiogram for routine diagnosis in the near future.

3. Vectorcardiography provides a rational explanation of conventional electrocardiography, and a knowledge of the vector principles aids interpretation of the electrocardiogram.

4. From the research point of view vectorcardiography has already added much to cardiac electrophysiology. When techniques of application are simplified and made more reliable the vectorcardiogram will surely extend basic knowledge in this field beyond the present limits which are set by electrocardiography. In both methods further advances require techniques of faster recording and greater amplification.

REFERENCES

Abildskov, J. A., Burch, G. E. & Cronvich, J. A. (1950) The validity of the equilateral tetrahedron as a spatial reference system. *Circulation*, **2**, 122.

Brigden, W. & Shillingford, J. (1952) The vectorcardiogram of cardiac infarction. *Brit. Heart J.*, **14**, 339.

Burch, G. E., Horan, L., Abildskov, J. A. & Cronvich, J. A. (1955) The spatial vectorcardiogram in posterior myocardial infarction. *Circulation*, **12**, 418.

Burch, G. E., Horan, L. & Cronvich, J. A. (1956) The spatial vectorcardiogram in anterior myocardial infarction. *Circulation*, **13**, 360.

Burch, G. E., Horan, L. G., Ziskind, J. & Cronvich, J. A. (1958) A correlative study of postmortem, electrocardiographic, and spatial vectorcardiographic data in myocardial infarction. *Circulation*, **18**, 325.

Burger, H. C. & Van milaan, J. B. (1946) Heart vector and lead, I. *Brit. Heart J.*, **8**, 157.

Burger, H. C. & van Milaan, J. B. (1947) Heart vector and leads, II. *Brit. Heart J.*, **9**, 154.

Burger, H. C. & van Milaan, J. B. (1948) Heart vector and lead, III. *Brit. Heart J.*, **10**, 229.

Donzelot, E., Milovanovich, J. B. & Kaufmann, H. (1950) *Etudes Pratiques de vectographie*. L'Expansion Scientifique Francaise, Paris.

Duchosal, P. & Sulzer, R. (1949) *La vectorcardiographie*. S. Karger, Basle.

Einthoven, W. (1903) Die galvanometrische Registrirung des menschlichen Elektrokardiogramms. *Pflueger's Arch. f. d. ges. Physiol.*, **99**, 472.

EINTHOVEN, W., FAHR, G. & DE WAART, A. (1913) Ueber die Richtung und die manifeste Grösse der Potentialschwankungen im menschlichen Herzen und ueber den Einfluss der Herzlage auf die Form des Elektrokardiogramms. *Pflüger's Arch. f. d. ges. Physiol.*, **150**, 275.

EDDLESTON, J. M. & LOWE, I. E. (1954) The spatial vectorelectrocardiogram in right ventricular hypertrophy. *Australasian Annals of Medicine*, iii, 18.

EDDLESTON, J. M. & LOWE, I. E. (1953) The spatial vector-electrocardiogram in the left ventricular hypertrophy of hypertension. *Australasian Annals of Medicine*, ii, 22.

FOSTER, R. P. & MURRAY, R. H. (1954) The *QRS* complex deformity of myocardial infarction. *Amer. J. Med.*, **17**, 587.

GRISHMAN, A. & SCHERLIS, L. (1952) *Spatial Vectorcardiography*. W. B. Saunders & Co., Philadelphia and London.

GRISHMAN, A., BORUN, E. R. & JAFFE, H. L. (1951) Spatial vectorcardiography: Technique for the simultaneous recording of the frontal, sagittal, and horizontal projections, I. *Amer. Heart J.*, **41**, 483.

MACKAY, W. R. (1958) Vectorcardiography in the diagnosis of myocardial infarction. *Progress in Cardiovascular Diseases*, **1**, 175.

ROBERTSON, D. (1951) On consistency in convention of sign in electrocardiography and vectorcardiography. *Amer. Heart J.*, **42**, 912.

SCHAEFER, A. I. (1957) Posture, the manifest heart vector, and the volume conductor. *Amer. Heart J.*, **53**, 895.

SCHELLONG, F. (1936) Elektrographische Diagnostik der Herzmuskelerkrankungen. *Verhandl. d. deutsch. Gesellsch. f. Inn. Med.*, **48**, 288.

SCHERLIS, L. & GRISHMAN, A. (1951) Spatial vectorcardiography: Left bundle branch block and left ventricular hypertrophy, II. *Amer. Heart J.*, **41**, 494.

SCHMITT, O., LEVINE, R. & SIMONSON, E. (1953) Electrocardiographic mirror pattern studies. I. *Amer. Heart J.*, **45**, 416.

SHILLINGFORD, J. & BRIGDEN, W. (1951) The vectorcardiogram in 100 healthy subjects using a new drawing instrument. *Brit. Heart J.*, **13**, 233.

SHILLINGFORD, J. & BRIGDEN, W. (1954) The vectorcardiogram of mitral stenosis. *Brit. Heart J.*, **16**, 13.

WALLER, A. D. (1887) A demonstration on man of the electromotive changes accompanying the heart's beat. *J. Phys.*, **8**, 229.

WILSON, F. N. & JOHNSTON, F. D. (1938) The vectorcardiogram. *Amer. Heart J.*, **16**, 14.

THE CATHODE RAY OSCILLOGRAPH AND ITS APPLICATION TO ELECTROMYOGRAPHY

A. Nightingale

The history of medical and biological instrumentation right up to the present day shows that the requirements demanded by the clinician and the research worker have always been at least as stringent, if not more stringent, than those met in other fields. The reader will no doubt be familiar with the researches of A. V. Hill into the heat developed during muscular and nervous activity, researches which led him to construct galvanometers whose performance can actually reach the limit of what is theoretically possible, in sensitivity and speed of response. The limit of sensitivity in this case is set by the Brownian motion of the galvanometer coil, which causes random fluctuations masking the signal. In this paper I have referred to a research of my own in which a similar limit had to be approached.

Similarly in electrophysiology also we find physiologists setting the pace in the application of new techniques and the development of old ones. By the early 1920's the Einthoven string galvanometer and the Lippman capillary electrometer had been developed and used to the utmost, and it had become clear that further progress in the measurement of the action potentials of nerves and muscles could only be made by instruments with a faster response and greater sensitivity. At this stage the valve amplifier and cathode ray tube came on the scene. The latter had of course originated at the end of the last century in the work on the electron by J. J. Thompson and others, notably Braun in Germany. The cathode ray tube had the required property of extremely rapid response, but was insensitive. The required sensitivity was provided by the valve amplifier. The first published cathode ray tube recording of an action potential, made by Gasser & Erlanger, in the USA, appeared in 1922. In this country physiologists continued to use the capillary electrometer in conjunction with valve amplifiers, and in 1928 the electromagnetic oscillograph was given a new lease of life through the development by Matthews of a rugged and reliable instrument with a more rapid response.

The main reason why this could still compete with the cathode ray tube was that the focusing and brightness of the trace on the cathode ray tube screen were so poor that a photographic record could only be made if the same wave-form was traced many times. That is, the cathode ray tube could only be used for repeated phenomena, not for a single transient event. In the Matthews oscillograph a beam of light was reflected from a mirror on to a photographic plate and records could be made of unrepeated events. However, in the early 1930's improvements in the focusing of the electron beam and in the fluorescent screen used in cathode ray tubes led to a much more brilliant trace and photographs of transient events became possible. These developments were no doubt accelerated by the use of the cathode ray tube in radio echo equipment to measure the height of the ionosphere, a technique which developed into radiolocation and radar. This progress led to the almost universal adoption of the cathode ray oscillograph for electrophysiological investigations, and since the Second World War the instrument has become widely used in all branches of biological and medical research and practice.

Section 1. The Cathode Ray Oscillograph

THE CATHODE RAY TUBE

A cathode ray tube is shown diagrammatically in Fig. 1. The envelope is made of glass and evacuated. A filament, through which an electric current is passed, heats the oxide-coated cathode *C*. This emits electrons

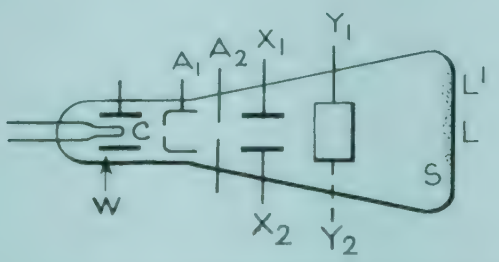


Figure 1. Cathode ray tube (electrostatically focused). Plan view.

which tend to spread out but are focused by the negatively charged cylinder *W* so that they pass through a small hole in the first anode A_1. A_1 is at, say, 1,000 volts and the second anode A_2 at 2,000 volts positive to the cathode. The electrons are accelerated along the axis of the tube and at the same time they are focused by the field between A_1 and A_2 on

to the fluorescent screen S. At the point where they strike the screen a bright spot of light appears.

Between the second anode and the screen are placed two pairs of deflector plates, X_1, X_2 and Y_1, Y_2. If a p.d. is applied between X_1 and X_2, with X_1 positive, the stream of electrons will be deflected towards X_1, and the spot of light will move from the centre of the screen L towards L^1. The diagram is a plan view, so that a deflection from L to L^1 represents a horizontal displacement. Similarly a p.d. between Y_1 and Y_2 produces a vertical displacement of the spot of light.

If a sinusoidal p.d. is applied between Y_1 and Y_2 the spot will move up and down as Y_1 is first positive and then negative with respect to Y_2. If the movement is rapid—more than a few cycles per second—a line is seen on the screen, because of the afterglow of the screen and persistence of vision of the eye. Now if a steadily increasing p.d. is applied between the X-plates the spot moves along horizontally at the same time as it moves up and down. The result is that the sinusoidal wave-form appears on the screen. What you really see is a graph of the p.d. applied to the Y-plates against the p.d. applied to the X-plates. If the latter increases steadily with time the picture is in effect a graph of p.d. against time. Thus the wave-form of an unknown alternating signal can be studied.

The great advantage of the cathode ray oscillograph as a measuring instrument over other forms of recorder (such as an electromagnetic pen writing on moving paper) is that the beam of electrons acts like an extremely light pointer, with very little inertia. It can thus follow very rapid changes in potential. In fact a cathode ray oscillograph can be used for investigating high frequencies of the order of 10 Mc./s., whereas a pen recorder can only be used up to about 100 c./s.

THE TIME BASE

This is the name given to the circuit which provides the time-scale. The time base must apply to the X-plates of the cathode ray tube a potential difference which rises steadily with time (Fig. 2). One traversal of the spot horizontally across the screen is called a 'sweep'. The sweep need not be right across the screen; sometimes one only uses part of the screen and the *sweep amplitude* is then small. Fig. 3 shows the effect of altering the sweep amplitude. The sweep amplitude is of course increased or decreased by increasing or decreasing the highest voltage reached by the time base.

The rate at which the spot moves horizontally is called the *sweep*

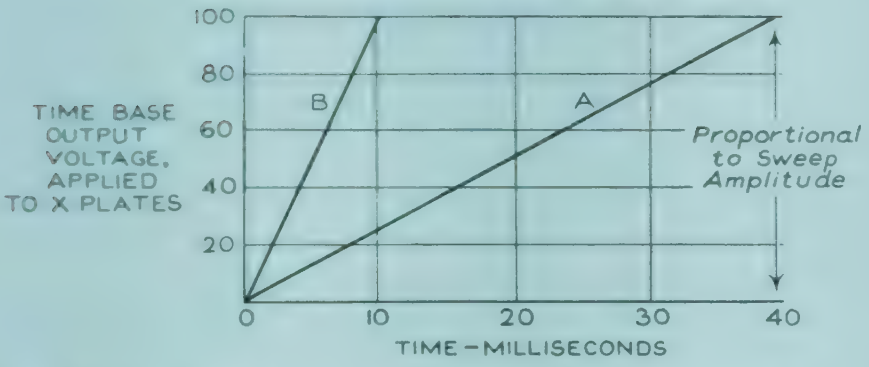


Figure 2. Showing how the time base voltage varies with time in order to draw the spot at a steady speed across the screen.

A: Relatively low sweep velocity.

B: Higher sweep velocity.

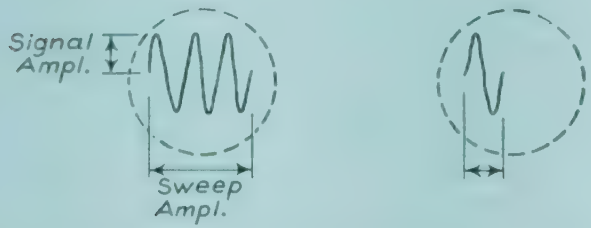


Figure 3. Effect of decreasing the sweep amplitude, when the sweep velocity is kept constant.

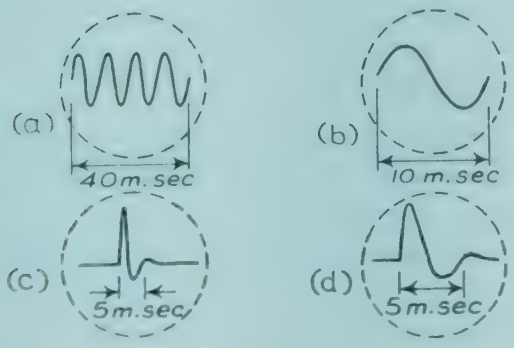


Figure 4. Effects of increasing the sweep velocity, when the sweep amplitude is kept constant.

a and *b*: Sinusoidal wave-form of frequency 100 c/s.

c and *d*: Action potential from a motor unit.

velocity. It is proportional to the rate at which the time base voltage rises, i.e. to the slope of the graph of Fig. 2. Line B has a greater slope and will give a greater sweep velocity than line A. Fig. 4 shows the effect of increasing the sweep velocity upon the appearance of the picture on the cathode ray tube. A higher sweep velocity implies a more open time-scale.

SINGLE SWEEP AND REPETITIVE TIME BASES

So far we have dealt only with a single traversal of the screen by the spot. At the end of the sweep the spot is at the right-hand side of the screen. In many cases one wishes to continue to examine the wave-form over a longer period of time, and to make this possible the spot must be brought rapidly back to the left-hand side of the screen (fly-back) in readiness for the next sweep, which should then commence immediately. The voltage between the *X*-plates must therefore fall very abruptly at the

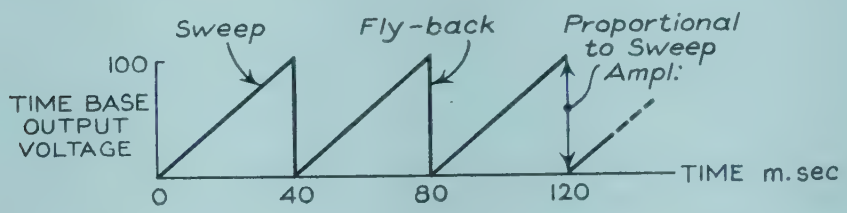


Figure 5. Repetitive time base wave-form.

end of the sweep, before repeating the same linear rise of voltage with time. The resulting 'saw-tooth' time base wave-form is shown in Fig. 5.

A time base giving only one sweep is called a 'single sweep' time base. One which repeats the sweep is called 'repetitive' or 'free-running'.

SYNCHRONIZED AND TRIGGERED TIME BASES

If a periodic wave-form is under observation it is usually found that the fly-back does not occur precisely at the end of a complete wave, so that each sweep draws out the wave-form in a position slightly displaced from the preceding one, and the picture drifts to the left or right. To avoid this it is possible to feed a proportion of the signal potential to the time base generator to ensure that the fly-back always occurs at the same point of signal wave-form and a steady picture is obtained. This process is called synchronization.

It is often convenient to 'trigger' the time base so that the sweep commences at a predetermined time. For example, this is done in the

measurement of conduction time in nerves. The electrical stimulus triggers the time base. Part of the stimulating pulse, and the action potential from the pick-up electrode on the nerve, both appear on the screen, the distance between them being a measure of the conduction time. The time base can be triggered from almost any part of a wave-form by means of suitable circuits, e.g. in an electrocardiogram the peak of the *R* wave or the point at which the *QR* excursion crosses the zero line.

LONG-PERSISTENCE TUBES AND METHODS OF RECORDING

The fluorescent coating of the cathode ray tube can be chosen so that the brightness of the trace decays either rapidly or slowly. A short persistence screen is convenient when a rapidly repeating time base is used, for examining a high frequency wave-form or phenomena of short duration such as action potentials. A long persistence of a few seconds in duration is useful when slower phenomena such as electrocardiograms are being observed, or for transient events. A recent development which has great potential value to medical and biological research is the 'infinite persistence' tube. This is a special type of cathode ray tube which 'stores' the picture. Once the trace has been written on the fluorescent screen it is retained for hours or weeks for study. It can be erased at will in a fraction of a second. A detailed account of storage tubes is given by Thomson & Callick (1959).

Permanent records from the conventional cathode ray tube are usually made photographically. To obtain a record of long duration the time base is switched off and the photographic film or paper in the camera moves continuously in a horizontal direction. The movement of the paper thus takes the place of the time base. Typical records of electromyograms taken by this technique are shown in Figs. 19–22. Alternatively, if the required information is contained in a single sweep of the time base, this can be photographed on stationary film (e.g. Fig. 16). The shutter of the camera is connected to electrical contacts so that when it is opened the time base is triggered. This ensures that only one sweep is obtained for each exposure. The principles and technique of photographic recording have been discussed by Hercock (1947).

Various methods are available for speeding up the processing of the records, and these can be very valuable. However, the photographic technique is costly and time-consuming, and in some of my own work I have recently returned to the earliest method of copying the picture directly on tracing paper held against the screen of the cathode ray tube.

Fig. 6 shows a record obtained in this way of the voltage and current wave-forms during stimulation of a dog's heart. This with Fig. 7 serve to illustrate the wealth of information which can be obtained from such a record; all three main components of the tissue impedance were evaluated. This information guided the design of an electronic pacemaker used for patients with atrio-ventricular block.

MULTIPLE BEAM SYSTEMS

It is often useful to present a number of different traces simultaneously, one above the other, on the cathode ray tube (for example, the electrocardiogram and the blood pressure). This can be achieved in several ways. One type of tube has a single gun and splits the beam of electrons into two parts after it has passed through the X-plates. The two beams are thus deflected by a common time base and in effect there are two separate pairs of Y deflection plates. Examples of records taken with this type of tube are given in Figs. 6, 18, 19 and 20. In another type of tube two or more separate electron guns are used giving complete independence. A technique which may be more economical in some circumstances employs a beam switching circuit. This system uses a normal single beam tube but switches the output from two amplifiers very rapidly between upper and lower positions on the screen, so that the effect of two separate traces is obtained. This system unfortunately cannot be used for high frequencies but is perfectly satisfactory for many biological applications.

THE COMPLETE CATHODE RAY OSCILLOGRAPH

The tube described above is electrostatically focused. Focusing and deflection by magnetic fields is also possible and widely used in television. However, most tubes commonly used in medical research employ electrostatic focusing. Since the p.d. required to deflect the spot across the screen is of the order of 30 volts per centimetre, it is usual to employ an amplifier with its output connected to the Y-plates in order to increase the sensitivity. The amplifier, the circuits for generating the time base wave-form, and the necessary power supplies, are all incorporated in a single cabinet. This constitutes the cathode ray oscillograph. A commercial unit now available provides facilities for measuring potential differences from 100 volts down to 10 millivolts and time intervals between several seconds and 1 microsecond. The input impedance of the vertical deflection amplifier is about 1 million ohms, so that we can regard the instrument as an efficient calibrated voltmeter and chronometer.

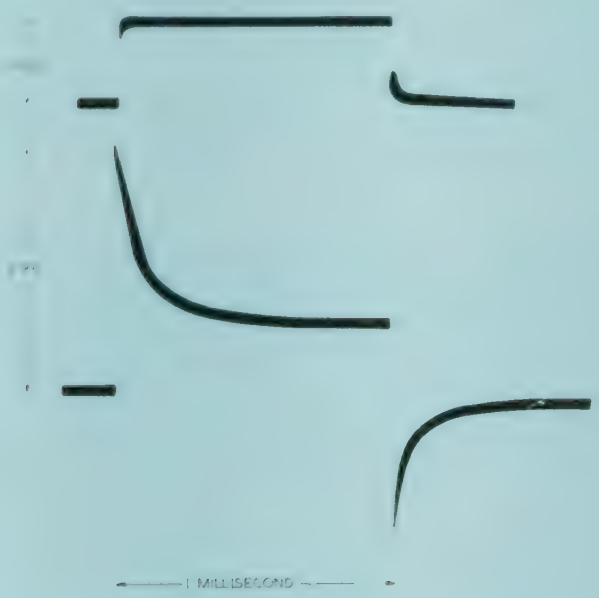


Figure 6. Tracing obtained from a double beam cathode ray tube, during electrical stimulation of a dog's ventricle.

Above: The voltage wave-form.

Below: The corresponding current wave-form. The 'spikes' at the beginning and end of the pulse show that the electrical properties of the tissues can be represented by an equivalent circuit containing a condenser as well as resistances (Fig. 7).

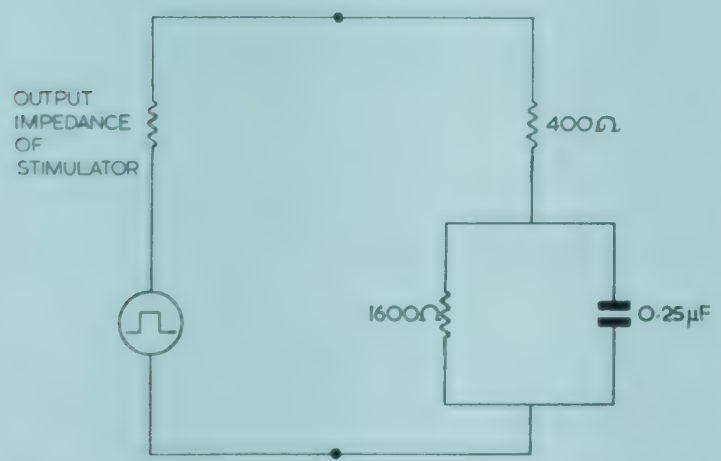


Figure 7. The equivalent circuit representing the tissues. The values were deduced by a detailed analysis of Fig. 6.

Another convenient system available commercially comprises a complete double beam oscilloscope having a short persistence tube and recording camera, with a second long persistence tube connected in parallel with the first. With the addition of pre-amplifiers designed for the particular purpose in hand, this combination can cope with many requirements encountered in medical research.

DISPLAY OF VARIABLES

The cathode ray oscillograph as described above is basically a device for drawing a graph of potential difference against time. Any quantity which can be converted into a potential difference can therefore be displayed.

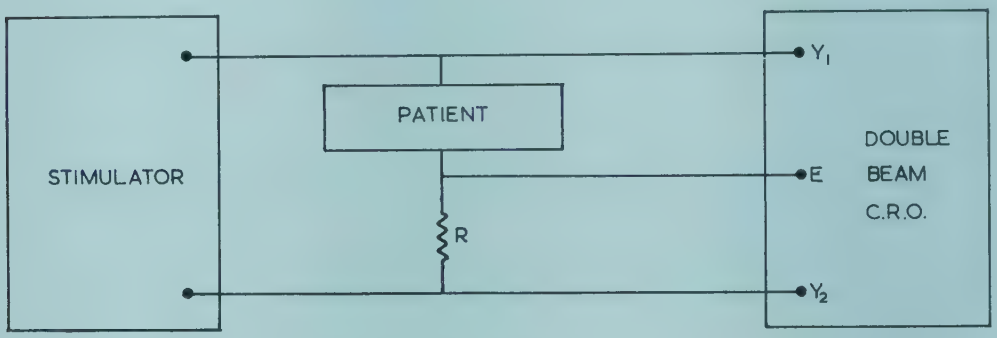


Figure 8. Circuit used to obtain both current and voltage wave-forms during electrical stimulation of a patient's muscle.

The simplest problem to be considered is how to measure a current. This is achieved by the simple trick of putting a small resistance (small enough not to disturb the current flow appreciably) in the circuit and using the CRO to measure the potential difference developed across the resistance. If we know the potential difference and the resistance we can calculate the current. Fig. 8 shows the circuit used for examining both current and voltage wave-forms simultaneously when stimulating a biological preparation or a human subject. This is in fact the circuit used to obtain Fig. 6.

Other quantities such as pressure, blood flow, etc. can be converted to electrical signals by means of suitable transducers. Furthermore we need not confine ourselves to a display of one variable against time. Any signal we like can be applied to the *X*-plates in place of the time base, so that in general we can plot a graph of any two variables. In vectorcardiography the two variables are the projections of the electrical vector

of the heart on two planes of reference. Transducers and vectorcardiographs have already been described in other papers of this series, and it will be sufficient here to mention one recent application of a novel method of display, used in research on the ultrasonic location of objects inside the body, a sort of ultrasonic radiography. In this technique a short pulse of ultrasonic radiation is transmitted into the tissues, and the

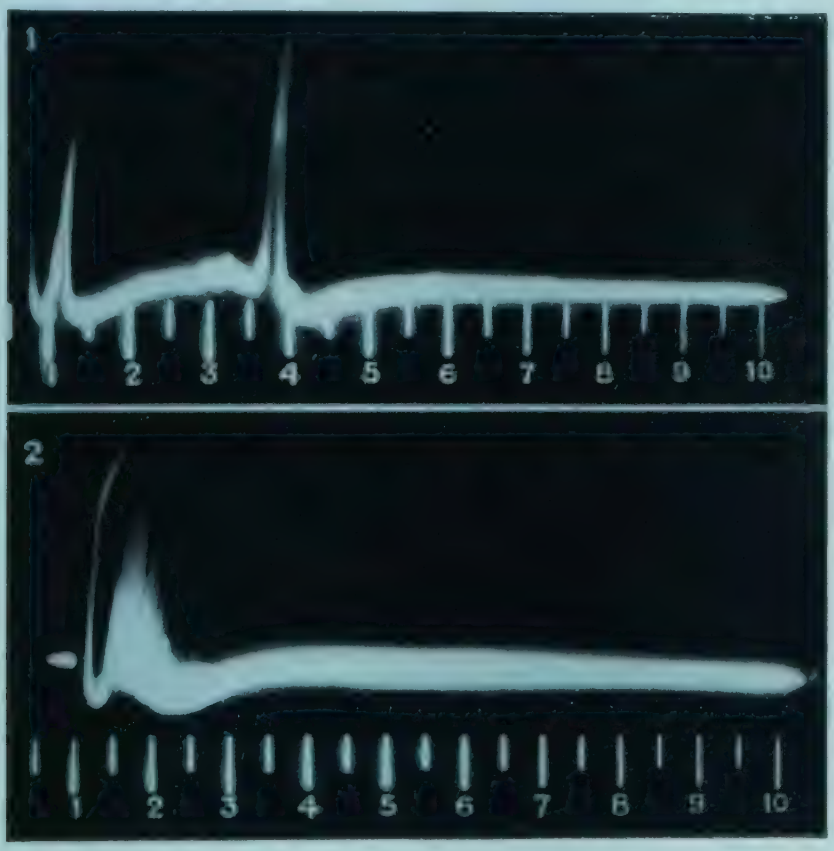


Figure 9. Ultrasonic 'echo sounding' technique; simple display of the transmitted and reflected pulses. The probe was placed over the bladder of a patient with acute retention of urine following colporrhaphy.
1: Distended bladder.
2: After the bladder had been emptied by means of a catheter.

echos from the various discontinuities inside the body are amplified. The simplest display is the standard one used in the non-destructive testing of welds and castings in industry. The transmitted pulse and the echo are both shown on the cathode ray tube, with a calibrated time base, and the time interval between the two is a measure of the depth of the discontinuity (Fig. 9). This gives information on a narrow cone of tissues only,

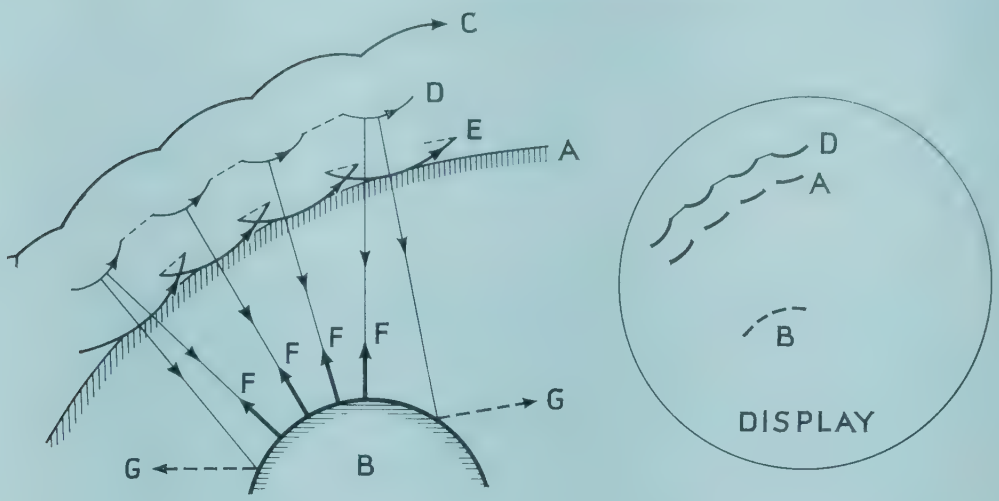


Figure 10. Method of obtaining a cross-sectional 'view' of the human thigh using ultrasonic pulses. A, patient's skin. B, reflecting mass. C, D, E, paths traced by the spindle in which the probe is mounted, by the transducers, and by the probe face respectively. F, paths of echoes returning to receiving transducer. G, paths of useless reflections. Pulses are sent out 50 times per second, and the probe is rocked at the same time as it is moved round the thigh.

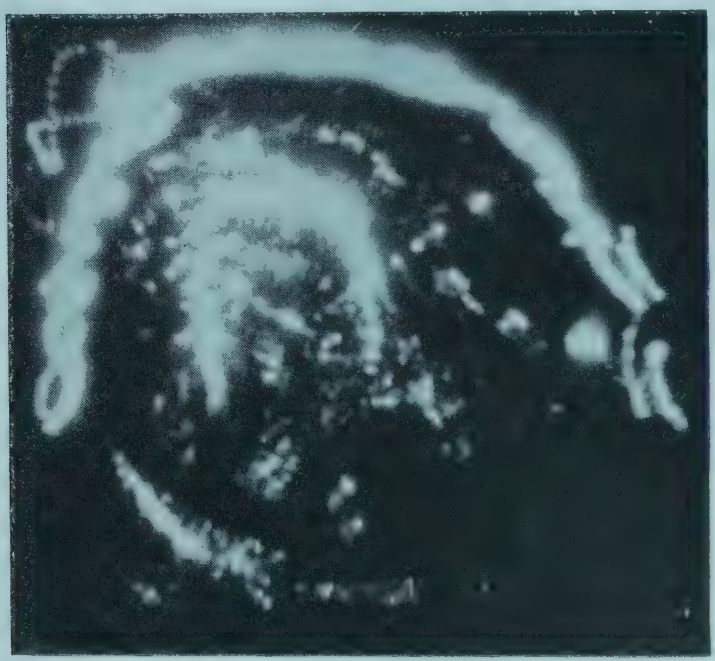


Figure 11. Cathode ray tube display obtained by the method explained in Fig. 10. The bright arc at the top represents the various positions of the probe as it was moved round the anterior aspect of the thigh. The bright area near the centre represents the femur.

the area covered by the transmitted pulse, and furthermore the range is very limited because of the rapid attenuation of ultrasound in tissue. A more sophisticated method of display, developed by Donald, MacVicar & Brown (1958), is shown in Figs. 10 and 11. Fig. 10 shows how the ultrasonic transmitter-receiver probe is moved round the thigh of a patient. Information on its position and angle of inclination are fed to the cathode ray tube in such a way that the point of origin of the sweep of the time base represents the instantaneous position of the probe, and the direction of traverse of the spot across the screen corresponds to the inclination of the probe. The brightness of the spot is modulated by the transmitted pulse and by the echoes, so that bright patches appear on the screen at points representing discontinuities in the tissues. As the probe is moved round, a picture like that of Fig. 11 gradually builds up on the long persistence screen.

Section 2. Electromyography

Electromyography is discussed here in more detail to illustrate some of the problems encountered in the use of the cathode ray tube and its associated equipment for a particular purpose.

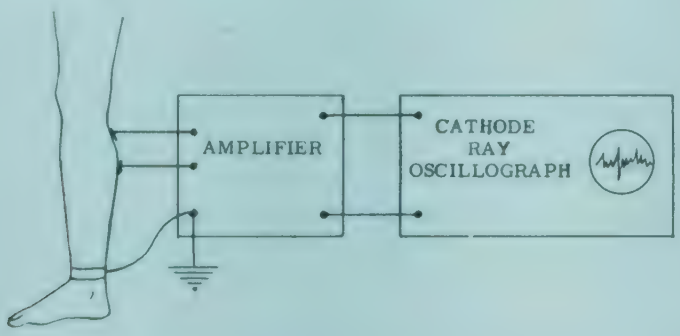


Figure 12. Block diagram of the apparatus required for electromyography.

The apparatus is shown in outline in Fig. 12. We can use surface electrodes, small metal discs stuck or pressed on to the skin of the subject, or a needle electrode inserted into the muscle under test. The type used in clinical electromyography is shown in Fig. 13. If the subject is asked to make a moderate or strong effort to contract his muscle, the electromyogram shows an irregular but continuous wave-form. However, if a very slight voluntary effort is called for, it is possible to detect the action potentials of a single motor unit (Fig. 14(*a*)). In a denervated muscle spontaneous fibrillation potentials attributed to single muscle fibres can

be seen (Fig. 14(*b*)). Fig. 15 is a schematic drawing of a motor unit, which in most skeletal muscles contains a hundred or more muscle fibres innervated by a single motor nerve fibre. Normally all these muscle fibres 'fire' almost simultaneously as an impulse arrives from the nerve, giving the relatively larger and longer duration motor unit potential of

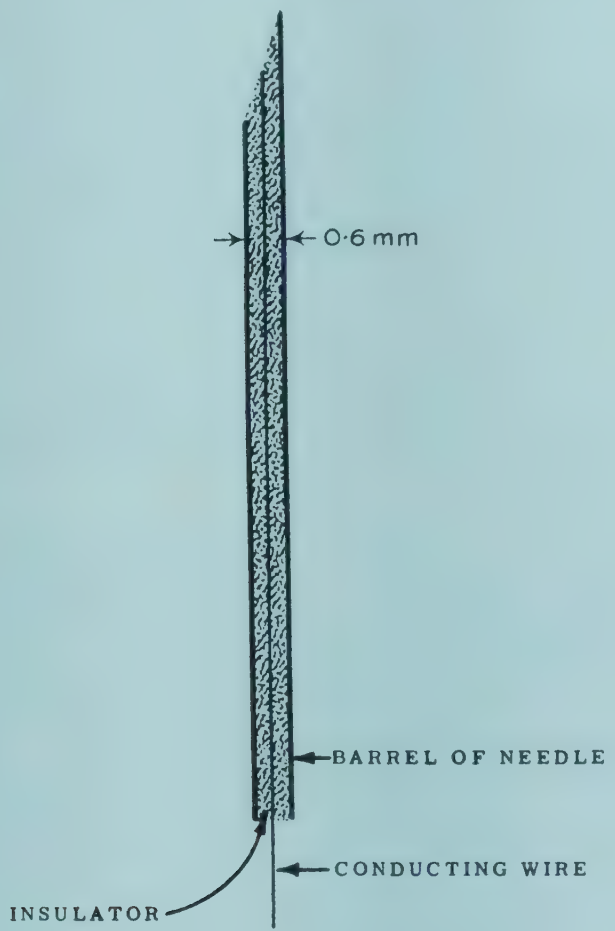


Figure 13. Concentric needle electrode.

Fig. 14(*a*). In fibrillation only one fibre contributes its individual action potentials, which are smaller and of shorter duration (Fig. 14(*b*)). Other abnormal wave-forms are shown in Fig. 16. A more detailed description of the physical principles has been given by Nightingale (1958b, 1959a) The clinical applications have been discussed by Richardson (1952), Rodriquez & Oester (1956) and Marinacci (1958).

We must not forget the value of quite different forms of presentation;

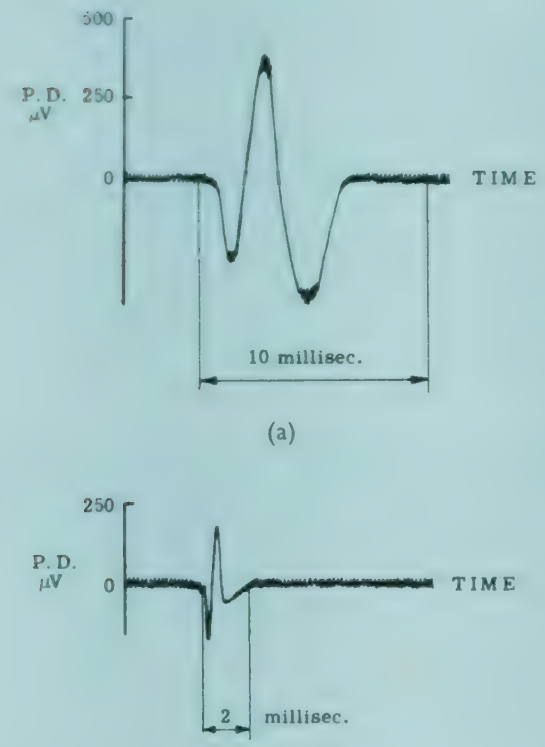


Figure 14. Typical wave-forms of action potentials picked up by a needle electrode (*a*) Motor unit. (*b*) Single fibre.

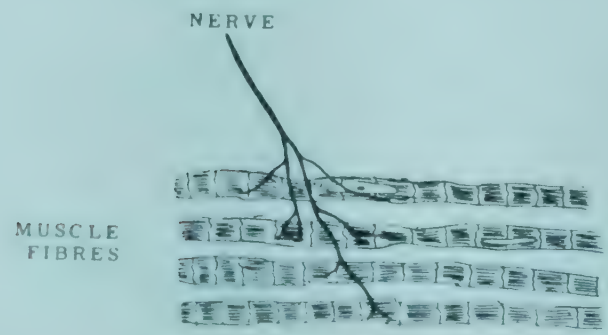


Figure 15. Part of a motor unit.

In clinical electromyography it is often found that the sound, produced by feeding the electrical signal to a loudspeaker, provides valuable information. For example, a normal motor unit action potential makes a 'thump, thump' sound, a fibrillation potential a 'tick, tick', and the

Figure 16. Recordings of electromyograms.
(*a*) Normal motor unit action potential, associated with a very weak muscular effort.
(*b*) An abnormal motor unit potential; the regeneration of motor nerves after damage.
(*c*) Single fibre potential; spontaneous activity in a denervated muscle.
(*d*) A burst of potentials; myotonia.

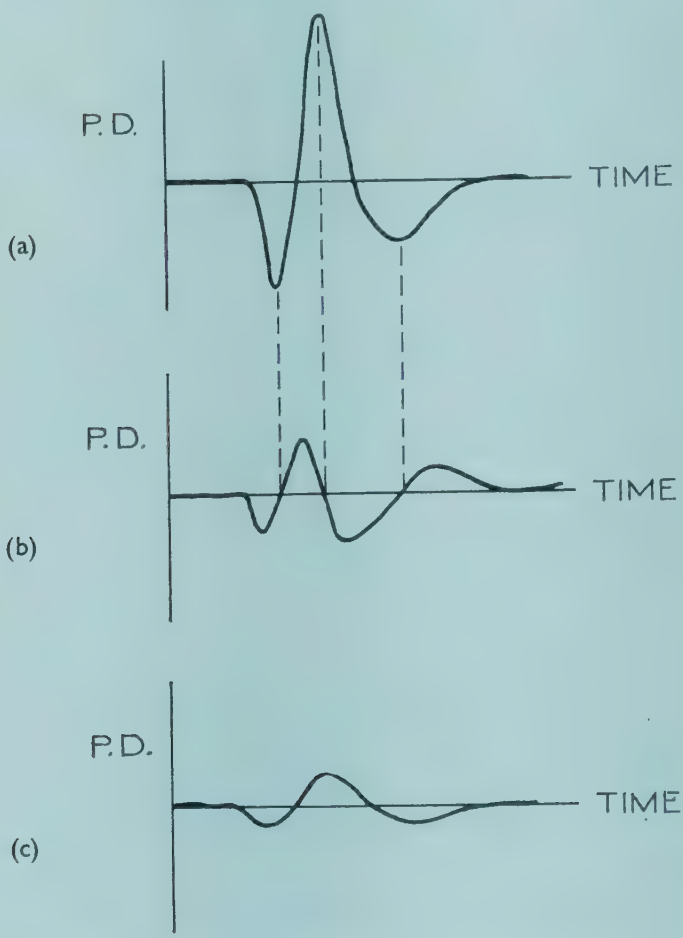


Figure 17. Distortion caused by amplifier.
(*a*) Motor unit action potential, faithfully reproduced.
(*b*) As recorded using an amplifier with inadequate low frequency response. The effect is analogous to the mathematical operation of differentiation.
(*c*) Inadequate high frequency response. In general there is a reduction in amplitude. In the extreme case integration occurs and only one excursion is obtained, representing the integral of the whole action potential.

sound of the burst of repeated potentials shown in Fig. 16(*d*), found in muscles in myotonia, has been likened by some who were in the RAF during the war, to the noise made by a dive bomber (e.g. Richardson, 1952).

These potentials are transient by their very nature and the clinician often needs to study them for a minute or so without the delay involved in making a permanent photographic record. This is achieved by a 'memory' unit. A widely used system contains a loop of magnetic tape which takes say 3 seconds for a complete revolution. The signal is fed to this tape continuously, and erased from it at the end of each revolution, so that at any instant the tape contains a record of any events which have

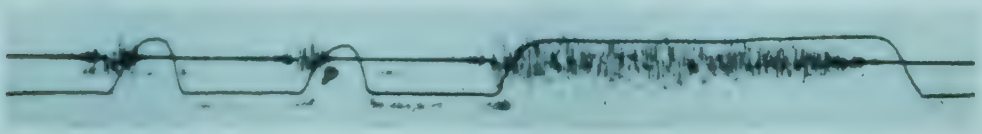


Figure 18. Recording from a strain gauge embedded in a molar of a person chewing a biscuit (two chews and a swallow are shown). The electromyogram was taken from surface electrodes over the jaw muscles. Note that the electrical activity commences before the mechanical effect. Interval between chews approximately 0·8 second.

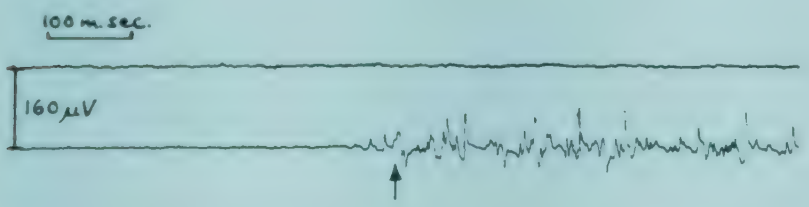


Figure 19. Recordings from electrodes over the thigh muscles of a subject in the stand-at-ease position. Upper: quadriceps femoris. Lower: hamstrings. During the recording the subject was asked to raise his arms forward. The arrow marks the point at which the arms had reached an angle of about 15° of flexion.

taken place during the preceding 3 seconds. A foot-switch enables the operator to stop the process of erase and record when he sees or hears an event of interest. The tape then plays back on to a cathode ray oscillograph whose time base is triggered by a reference point on the tape. Any parts of the 3-second record can then be selected and expanded at a greater sweep velocity so that the detailed wave-form can be studied. An improved system developed at St Thomas's Hospital employs a magnetic sheet revolving on a drum, so that a large number of records can be taken side by side as the recording head is moved along the drum.

For faithful recording of these electromyograms we must be able to

amplify a sufficiently wide range of frequencies. Analysis of the action potential wave-form shows that it contains components in the range 2 c./s. to 20 kc./s. (Buchthal, Guld & Rosenfalck, 1954). If the band of the amplifier is inadequate, distortion results (Fig. 17). For clinical electromyography it is usually sufficient to take a band between 50 c./s. and 3 kc./s. (Buller & Styles, 1952).

Apart from its clinical applications, electromyography is of value in the investigation of the activity of the muscles of the normal subject in various positions or movements (e.g. Åkerblom, 1948; Floyd & Roberts, 1958; Joseph, 1959). For this purpose it is usual to employ surface electrodes rather than needle electrodes if possible for obvious reasons. Fig. 18 shows

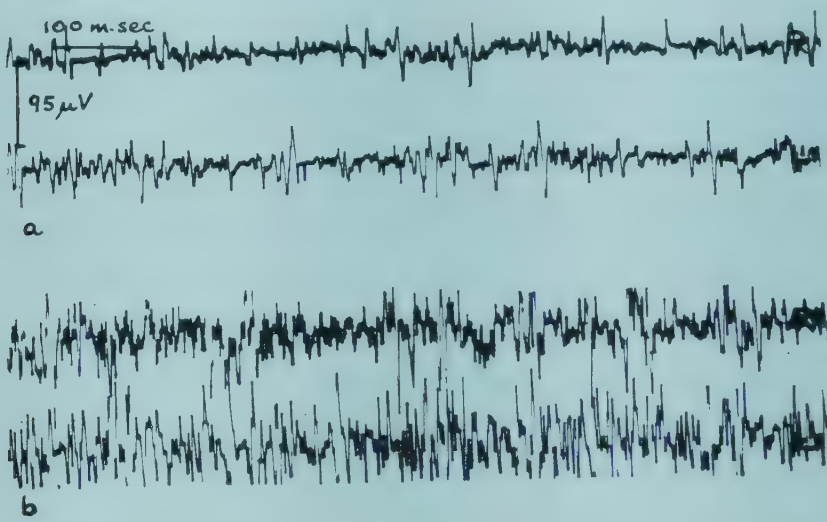


Figure 20. Recordings from electrodes over the left and right soleus muscles, for a subject in the stand-at-ease position.
(*a*) Wearing low-heeled shoes.
(*b*) Wearing high-heeled shoes.

two simultaneous recordings on a double beam CRO of the force exerted by the teeth in chewing and the corresponding electromyogram from the jaw muscles. In Fig. 19 the two records are taken from the front and back muscles of the thigh in the stand at ease position. At the beginning of the record we see that there was no detectable activity in either muscle, indicating that in this position the knee can be 'locked' without the need for muscular effort. During the recording the subject raised his arms forward. This shifted his line of weight so far in front of the knee joint that the back thigh muscles became active (Joseph, 1959). Fig. 20 shows that wearing high heels is accompanied by an increase in the activity of the calf muscles in the stand-at-ease position (Joseph & Nightingale, 1956).

The variation of electrical activity from muscles in the forearm during the action of grasping an object such as a screwdriver is shown in Fig. 21. Initiation of the movement gives a large burst of signals, followed by a lower steady level as the effort is maintained. This particular recording was taken during a study of the possibility of using the signals, from the stump of a patient whose hand had been amputated, to control the operation of a prosthesis. This investigation was initiated several years ago by the late Professor Whillis (Battye, Nightingale & Whillis, 1955). We showed that the technique was possible in principle. We made no attempt to miniaturize the equipment, but recent developments in the

Figure 21. Recording from electrodes placed on the forearm of a subject who was asked to grip and hold a screwdriver.

applications of transistors, small electric motors and pneumatic 'muscles' (Geddes, Moore, Spencer & Hoff, 1959) have made it possible to envisage a practical equipment. One feature of Fig. 21 is of particular interest. This is the momentary cessation of activity, lasting some 40 m./sec., between the initial movement and the sustained contraction. It is possible that this is caused by a momentary inhibition of the muscle by signals from the pressure sensitive organs of the fingers. This phenomenon could be used in time and motion studies where a detailed analysis is required of the various stages of a movement.

In this type of work we are often trying to detect a small electrical

signal level associated with a slight muscular contraction, and this may be masked by interference and electrical fluctuations generated in the amplifier itself. In fact we found it necessary to go to considerable lengths to reduce this amplifier noise, and we eventually achieved a noise level lower than any previously reported in the literature (Nightingale, 1958a). With this apparatus we then found it possible to detect an electrical fluctuation from a muscle and also from non-muscular sites such as the tibia (Fig. 22), even when the subject was resting comfortably on a bed (Joseph, Nightingale & Williams, 1955). The origin of this 'background noise' from tissues is not yet established (Nightingale, 1958b, 1959b), but

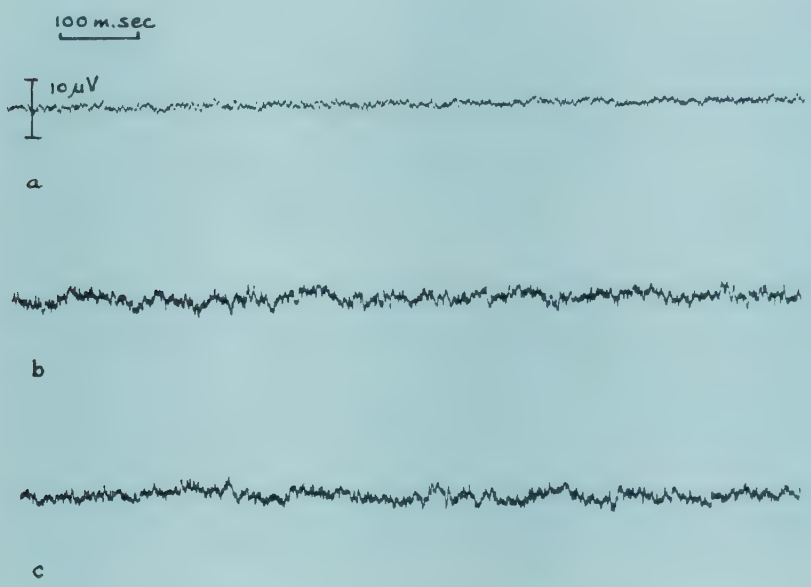


Figure 22. Illustrating the 'background noise' level during electromyography with surface electrodes 1 cm. in diameter.
(*a*) Amplifier noise.
(*b*) Electrodes over a relaxed muscle.
(*c*) Electrodes over the tibia.
The subject was lying comfortably on a bed.

it does represent the true base-line in electromyographic investigations. It also sets a limit to the necessary sensitivity of the amplifier, since a greater amplification than that used for Fig. 22 simply enlarges the 'noise' without adding any useful information.

When the electromyogram, taken with surface electrodes, is analysed it is found that the important components lie in a rather lower frequency range than is the case for clinical electromyography, in which the shape

of the individual motor unit potential is important. An analysis for different degrees of voluntary effort is shown in Fig. 23 (Nightingale, 1954*b*). For faithful reproduction, to ensure that nothing significant is missed, and for accurate quantitative measurements the amplifier must cover a frequency range between about 10 c. s. and 1,000 c. s. However, for comparative studies at fairly high levels of muscular effort a much smaller range is adequate. Floyd & Silver (1952) showed that a pen recorder could be used. Fig. 24 shows two records taken from the same pair of surface electrodes. The lower trace is the output from a wide-band

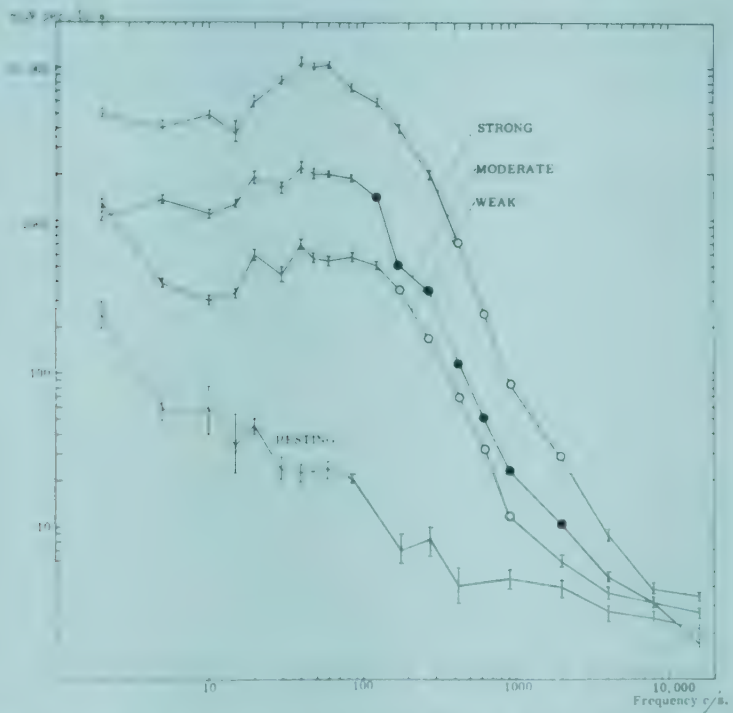


Figure 23. Spectra of electromyograms taken from the soleus muscle of a subject at rest and then making a 'weak', 'moderate' and 'strong' effort respectively. The curve marked 'moderate' corresponds to the normal standing position. The ordinate was obtained by taking the square root of the mean square voltage per unit bandwidth, using a Muirhead-Pametrada Wave Analyser.

amplifier, and the upper is the output from an amplifier whose gain has been reduced at the higher frequencies so that it corresponds roughly to that of a pen recorder. It is clear that there has not been much loss of useful qualitative information.

This account of electromyography illustrates the types of problem normally encountered in any application of the cathode ray oscillograph. Firstly the type of information required is considered and the frequency

range necessary for its faithful presentation. Next the required sensitivity is estimated, and it may be necessary to approach the theoretical limits of amplification. Details of the technique can then be settled, the method of taking permanent records, whether a long or short persistent screen is required, and whether a memory system is desirable.

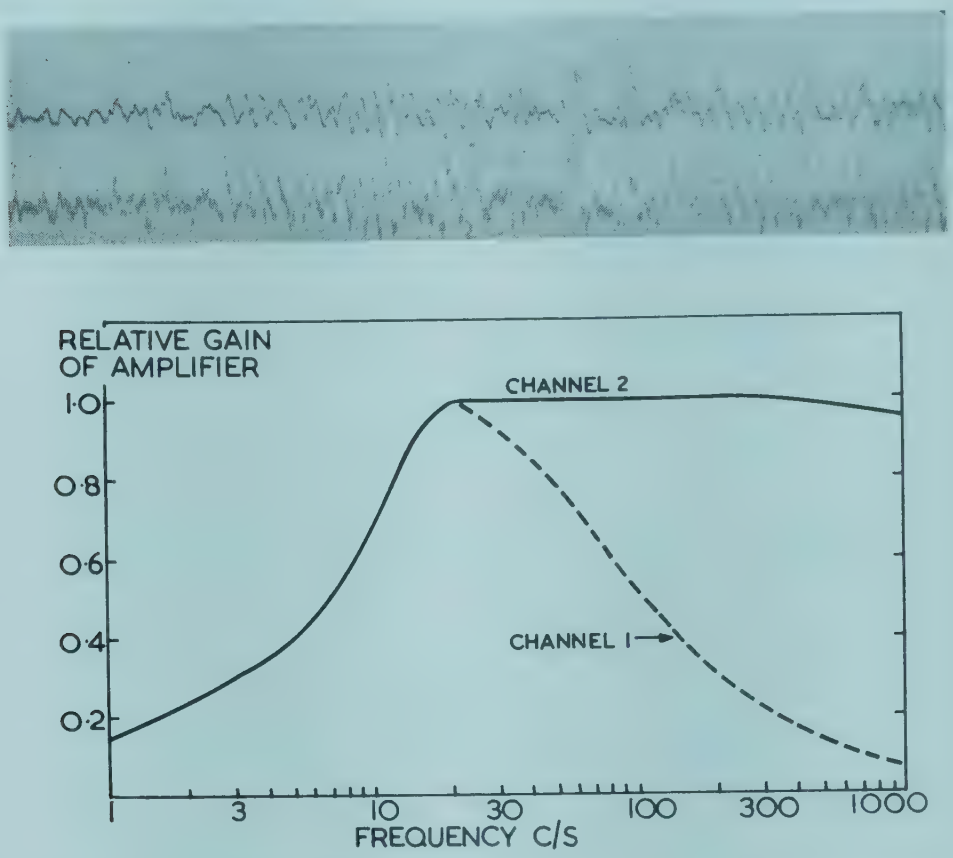


Figure 24. Showing the effect of reducing the high frequency response of the amplifier, upon the quality of the electromyogram. Surface electrodes. Simultaneous recordings from the same electrodes. Upper trace (channel 1): using a restricted band of frequencies. Lower trace (channel 2): using a wider band of frequencies.

COMPARISON WITH OTHER INSTRUMENTS

One great advantage of the cathode ray oscilloscope over other forms of recording instrument is its tremendous speed of response. The electron beam has a very small inertia and follows the impressed changes in potential extremely rapidly. In practice the response time is often limited by the quality of the amplifier. A typical commercial CRO will display without distortion alternating potential differences between d.c. and 5

million cycles second. It will respond to a sudden change of potential in less than 0·1 millionths of a second.

For comparison let us look for a moment at alternative types of instrument. The typical ink-writing pen galvanometer used in electrocardiography records up to 100 c. s., its response being limited by the inertia of the moving system and the friction of the pen against the moving paper. An instrument is now available in which the moving stylus is replaced by a fine jet of ink. This takes up to several hundred cycles per second. If we are content to obtain only a photographic record, with the delay involved in processing, we can use a beam of light reflected from a mirror as the pointer and obtain a frequency response up to several thousand cycles per second. The processing difficulty has been ingeniously overcome in a recent design in which ultra-violet light is used, recording on a special ultra-violet sensitive paper. This gives a visible record immediately. These recorders all have their place in medical instruments, but none can compete with the cathode ray oscillograph in speed of response. As we have seen, it is important in clinical electromyography to record faithfully frequencies up to several thousands of cycles per second, and the cathode ray oscillograph remains the instrument of choice in this field.

ACKNOWLEDGMENTS

I should like to thank the Medical Illustration Department, Guy's Hospital Medical School, for help in the preparation of the diagrams. I am grateful to the following for permission to reproduce figures: Dr J. Joseph and Charles C. Thomas, Publisher, Springfield, Illinois (Figs. 19, 20 and 22); Dr D. J. Anderson (Fig. 18); Mr C. K. Battye and the Editor, *Journal of Bone and Joint Surgery* (Fig. 21); Professor Donald and his co-workers, and the Editor, *Lancet* (Figs. 9, 10, 11); Messrs G. Bell and Sons Ltd (Figs. 1, 2, 3, 4, 5, 16, 17).

REFERENCES

Åkerblom, B. (1948) *Standing and Sitting Posture*. A.-B. Nordiska Bokhandeln, Stockholm.

Battye, C. K., Nightingale, A. & Whillis, J. (1955) The use of myo-electric currents in the operation of prostheses. *J. Bone Jt. Surg.*, **37B**, 506–510.

Buchthal, F., Guld, C. & Rosenfalck, P. (1954) Action potential parameters in normal human muscles and their dependence upon physical variables. *Acta physiol. Scand.*, **32**, 200–218.

Butler, A. J. & Styles, P. (1952) Some technical factors affecting the accuracy of recording in electromyographic investigations. *Annals of Physical Medicine*, **1**, 37–49.

Donald, I., MacVicar, J. & Brown, T. G. (1958) Investigation of abdominal masses by pulsed ultrasound. *Lancet*, June 7th, 1958, 1188–1195.

Floyd, W. F. & Silver, P. H. S. (1952) Comparative studies of ink-writer and cathode ray oscilloscope for electromyography. *J. Physiol.*, **117**, 36P.

Monograph No. 1. Ilford, London.

MICRORADIOGRAPHY

[illegible]

The extent to which X-rays can penetrate material is related directly to the voltage used to drive protective and intensify at their wave-lengths. In medicine those penetrate from a higher voltage and yet short wave-length, the so-called 'hard' X-rays have been employed for both diagnosis

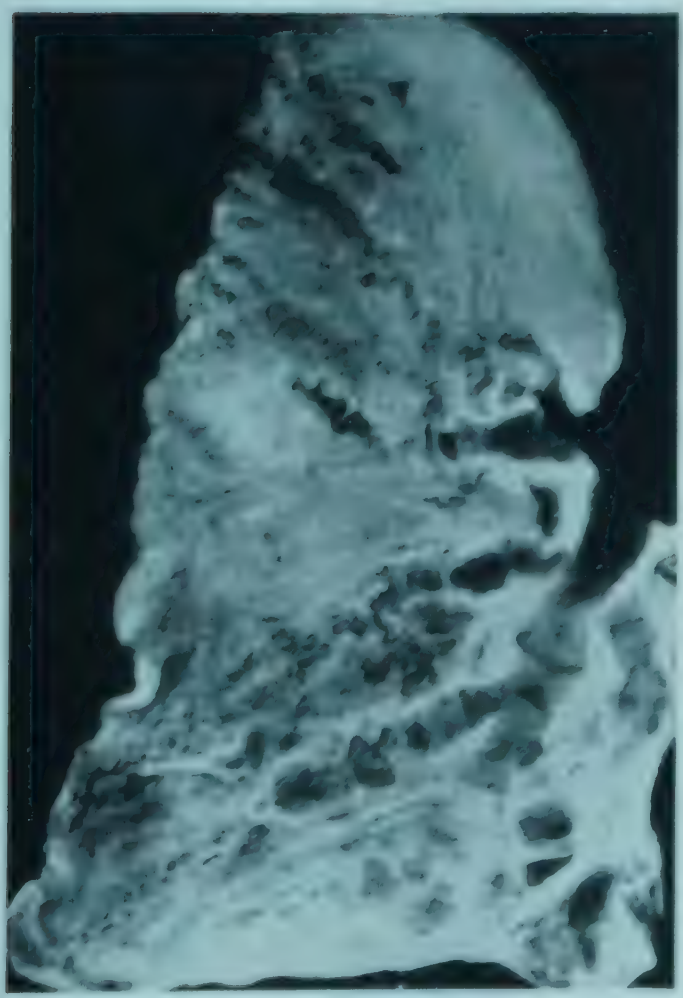


Figure 1. [illegible] radiograph of [illegible] lung showing [illegible] [illegible] prepared at [illegible] and [illegible]

and treatment. The X-rays first discovered in 1895 by Röntgen were hard in character and on this account passed through soft tissues, being only obstructed by bone. Thus, when Röntgen observed a radiograph of the bones of his own hand the main line of application of X-rays to medicine was indicated. Later, the introduction of radiopaque materials was used to visualize certain soft tissue structures and the armamentarium of the diagnostic radiologist increased. In all these examinations X-rays generated from a kilovoltage of 20–120 were employed. With the introduction of radio-therapy harder X-rays were used and these were generated from voltages of 250–1,000 kV. or even more. It thus happened that soft X-rays generated from a kilovoltage of 20 or less were neglected in medical circles.

Some years ago, in the course of a research on pulmonary conditions using the injection of radiopaque materials, it was accidentally observed that remarkably good radiographs of soft tissue could be obtained without any prior injection. For this purpose X-rays generated at 15 kV. were employed and pictures of whole lobes of lung were obtained (Cunningham & Miller, 1952; Cunningham, 1955). In cases of bronchiectasis the dilated bronchi were easily visualized (Fig. 1).

The pathology of some pulmonary conditions was very clearly illustrated, especially if stereoscopic radiographs were taken and studied in a special viewer designed for the purpose. Lung was particularly well suited for examination as the air in the bronchial tree gave excellent radiographic contrast with soft tissue. The factors governing the clarity of such radiographs are best illustrated by Frank's formula (Bohatyrschuk, 1944).

$$W = \frac{b\,d}{F - b}$$

where W = blurring of image.

b = distance of object from film.
d = size of focal spot.
F = target film distance.

As the specimen was practically in contact with the film the important factors in reducing blurring are the smallness of the focal spot, and the distance of the object from the film. Of these two the size of the focal spot is of much greater importance and in the X-ray tube used for this study measured 0·3 mm. For producing a focal spot of this small size an ingenious device known as a line focus has been devised. Thick slices of tissue were radiographed—up to 0·5 cm. thus enhancing

the stereoscopic effect. About the same time similar studies of bone had been performed (Sissons, 1950). Although these examinations of bone and lung were encouraging they nevertheless only enabled 'low power' views of tissue to be studied. If examination was to proceed to a cellular or microscopic level it was apparent that a modification of this technique would be necessary.

Although soft X-rays had not been applied until recent years to biology and medicine, it is surprising to learn that they had been much used in other fields. One of the earliest studies of soft tissues was made by Goby (1913) who with limited facilities studied diatoms and foraminifera. The X-rays employed were not really soft by modern standards and the coarseness of the photographic emulsions then available prevented a magnification of more than 17 diameters being reached. In later years similar studies met with considerable success (Fricke, 1932; Sherwood, 1934, 1936, 1937). Insects and seeds were studied in addition to the distribution of radiopaque organic spray materials on foliage (Eastman Kodak, 1943). The first successful radiograph of a histological section of plant tissue was made by Dauvillier (1930), who claimed to have obtained good resolution at a magnification of 600 diameters. The future possibilities of these techniques were indicated by Lamarque (1938) who obtained beautiful radiographs of histological sections of skin and named the technique historadiography. The detail in these radiographs compared very favourably with that seen in photomicrographs taken with the light microscope. The importance of Lamarque's work lay in the fact that he employed very soft X-rays (generated at about 2 kV.). As these X-rays do not even penetrate air very well he was obliged to enclose the specimen and photographic plate in a vacuum chamber. The widespread use of the soft X-ray technique at the time can be appreciated when one hears that it was applied to wood for the detection of death-watch beetle (Seeman 1943), in the textile industry for studying the weave of cloth (Sherwood 1936) and even for the detection of forgeries in postage stamps (Cheavin 1947). Copies of paintings by famous masters have been detected by radiography because of a difference in the composition of the paints employed. An X-ray examination of the Reynolds' portrait of John Hunter in the Royal College of Surgeons was recently made at the instigation of Sir Reginald Watson-Jones. It was found that a previous portrait, painted probably some three years earlier, existed under the present one and represented Hunter before the effects of severe illness were apparent in his thinned-out features (Lefanu, 1959).

In this short historical survey mention must also be made of the pioneer

work in this country by Barclay who was able to employ slightly harder X-rays as his classical studies in microradiography were made following the injection of radiopaque materials (Barclay, 1947; 1949; 1951; Trueta *et al.*, 1947). Bone was studied by Sissons (1950) and soft tissues by others (Bohatyrschuk, 1944; Wolff, 1950; Tirman *et al.*, 1951). In more recent years the work of Engström *et al.* (1955) has done much to promote the active interest in these techniques at the present time. As better and better resolution was obtained so the use of the term historadiography tended to be discarded and replaced by X-ray microscopy. It now appears likely that even the finer structures of the cell can be studied and in some cases the chemical constitution of certain cellular elements can be calculated.

TECHNICAL

Examination of material by X-rays rather than by visible light may be advantageous for two reasons—

(1) Since the resolution of an object is related to the wave-length of the examining medium, X-rays which have a wave-length about 1,000 times shorter than visible light rays should give much better resolution.

(2) Examination of an object in depth is possible owing to the great penetrating power of X-rays.

During the past fifteen years much research has been devoted to devising suitable techniques for X-ray microscopy. Two methods, known respectively as contact and projection microradiography are in use at the present time, whilst a third, namely reflection X-ray microscopy, is still in the stages of development. In all these techniques 'soft' X-rays have been employed. Although there is no recognized definition of 'hard' and 'soft' X-rays we shall adopt the following arbitrary classification.

Types of X-ray

Type	Wave-length	Kilovoltage
Ultra-hard (therapy)	less than 0·1 Å	Above 120
Hard (diagnosis)	0·1–1 Å	15–120
Soft	1–10 Å	5–15
Ultra-soft	More than 10 Å	less than 5

As the X-ray beam generated at a particular kilovoltage contains a range rather than a single wave-length, the above values are approximate.

CONTACT RADIOGRAPHY

This method is really an extension of ordinary radiography in which the specimen is in contact with the recording photographic emulsion. The image is therefore identical in size with the object and any magnification required is obtained photographically. The image must therefore be sharp and the photographic emulsion a very fine one. As with visible light, a small source is essential for a sharp image and so the focal spot is usually of the order of 0·3 mm. With the appearance of commercial fine-grain photographic emulsions contact microradiography has become a very simple and practical procedure, and it is for this reason that most of the work performed in the field has been done using contact methods. For this purpose soft and ultra-soft radiation have been used. The results obtained may be described as (*a*) qualitative, (*b*) quantitative.

The qualitative results include the demonstration of small blood-vessels following the introduction of contrast media, and much of Barclay's pioneer work was of this nature. The animal was injected during life, then sacrificed and sections made from the animal tissues and subsequently radiographed. Bellman (1953) devised a technique where the capillaries could be observed in the living animal, though examination could only be made in areas of the body such as the rabbit's ear where the tissue was not too thick. Similar studies were made of the vascular patterns in muscle and skin by Saunders *et al.* (1957). For these methods X-rays generated from 15 kV. could be used, whereas for examination of uninjected soft tissues a lower voltage was employed and microradiographs were obtained which bore a close resemblance to photomicrographs of the same tissue.

While Engström and his co-workers at the Karolinska Institute, Stockholm, played a great part in the development of these methods their contributions to the quantitative side of the subject are even more important. They demonstrated that it was possible to weigh a histological structure by estimating the attenuation of X-rays by means of densitometry in the microradiograph. This procedure can be even more valuable in identifying actual constituents. If it is possible to remove a substance from the specimen either by biochemical extraction or other means, it is possible to calculate the amount of the substance by difference (Engström & Lindström, 1950; Lindström, 1955). Using these methods Engström & Glick (1956) were able to estimate the weight of the chief cells of the gastric mucosa, and to demonstrate that dry weight of the nucleoli was high. For these estimations an ordinary beam of polychromatic X-rays is

used. A further extension of the quantitative method involves the use of a monochromatic beam. It is well known that every element shows a characteristic absorption curve in regard to X-rays of different wave-lengths (Fig. 2). This absorption curve contains a number of 'edges' where a sharp change occurs in the amount of absorption. By selecting monochromatic wave-lengths close to and on either side of an absorption edge of an element it is possible to calculate the amount that is present, by the difference in absorption at the two wave-lengths.

For the success of this method the substance must be present in sufficient concentration to give a measurable difference in absorption using two different wave-lengths. Furthermore, other substances with overlapping

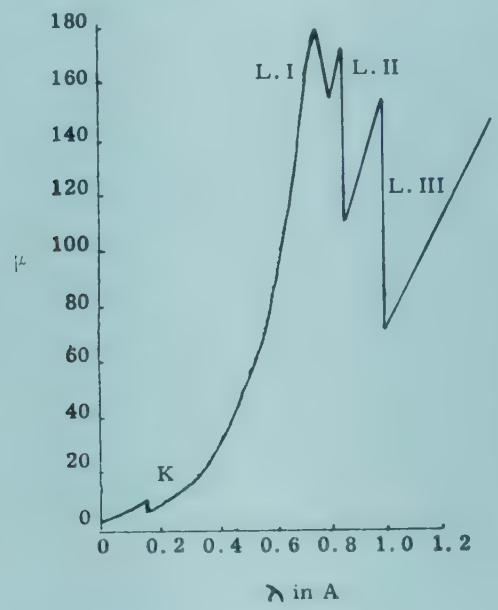


Figure 2. Showing the characteristic 'absorption edges' on an X-ray absorption curve.

ranges of absorption must of course not be present. It follows, therefore, that the method is unsuitable for substances present in low concentration. In spite of technical difficulties the amounts of such substances as carbon, nitrogen, oxygen and sulphur have been estimated, though better results have been obtained in the case of calcium and phosphorus.

THE SPECIAL CASE OF BONE

If undecalcified bone be used, the presence of radiopaque calcium salts permits the use of X-rays harder than those which must be used to

examine soft tissues. X-rays produced at 10–35 kV. are suitable and the apparatus becomes simpler. Vacuum cassettes are not necessary and the X-ray tube can be provided with a beryllium window 0·2–1 mm. thick.

Amprino & Engström (1952) have shown that even in normal bone, the bone salts are unevenly distributed. This work has been followed up in this country by Sissons, who has studied bone from a large number of anatomical sites by histological, chemical and microradiographic methods. Quantitative results have been obtained using an undecalcified section, the thickness of which is accurately known. The densities of various parts

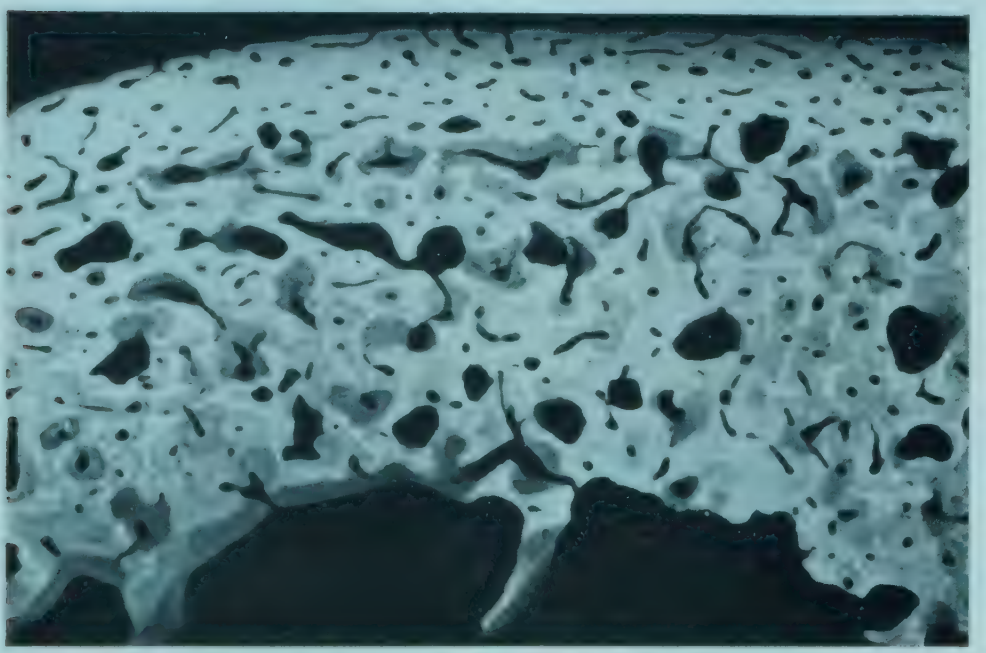


Figure 3. ×14. General view of cortical bone of mid-femoral shaft showing conspicuous bone formation and bone resorption. From a male aged 2½ years (by kind permission of Dr H. A. Sissons).

of the microradiograph are estimated microphotometrically by comparison with an aluminium reference system, and the amount of hydroxyapatite in any given area of a bone section precisely determined (Sissons, Jowsey & Stewart, 1959).

Using this technique it is possible to find out the state of activity of the bone surface which is being assessed. The areas of breakdown appear as large crenated spaces, whilst areas in which bone is being laid down appear greyish in contrast to the whiter areas of mature bone. These results have been proved to be accurate as they show good correlation

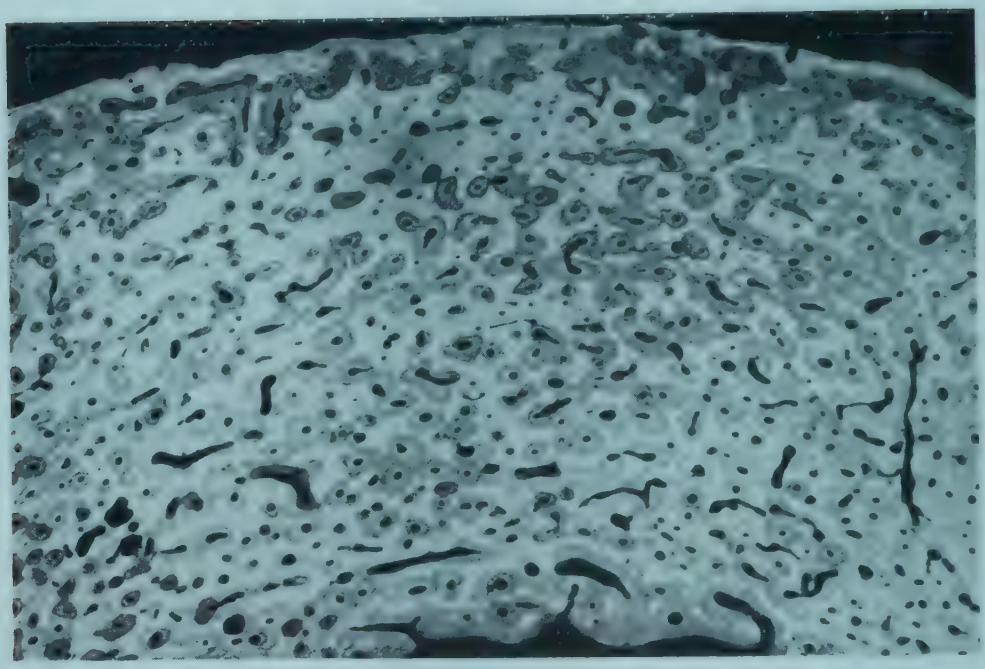


Figure 4. ×15. General view of cortical bone of mid-femoral shaft showing little evidence of formation or resorption of bone. From a male aged 20 years (by kind permission of Dr H. A. Sissons).

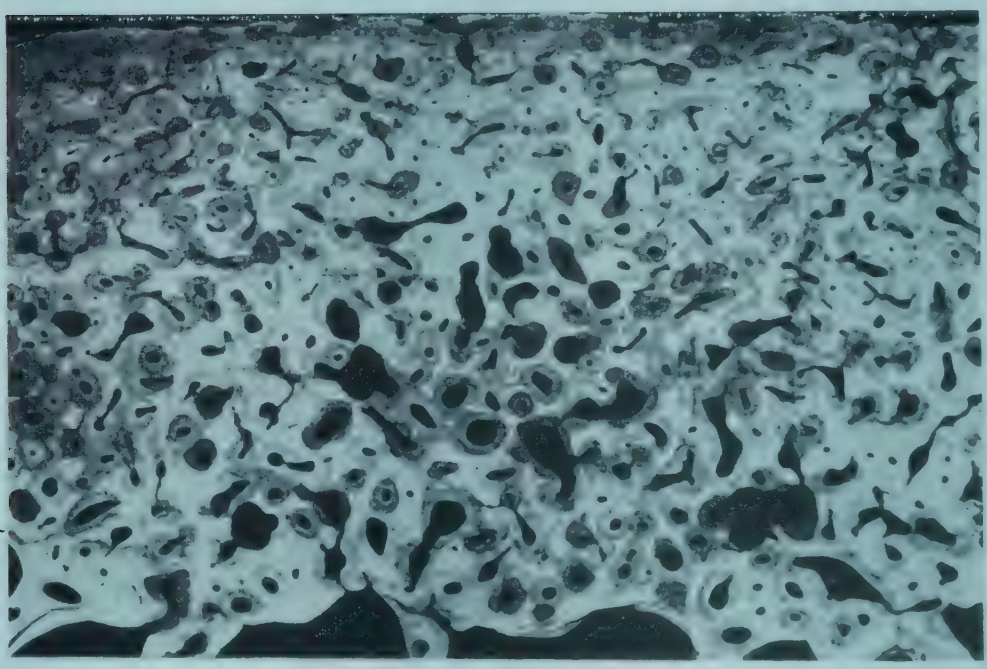


Figure 5. ×14. General view of cortical bone of mid-femoral shaft showing increased bone resorption and reduced bone formation. From a male aged 65 years (by kind permission of Dr H. A. Sissons).

with histological studies and with uptakes estimated by isotopic techniques. It has been found that in bones from children (Fig. 3) there is conspicuous evidence of both formation and resorption of bone; in a young adult (Fig. 4) there is little evidence of either formation or resorption of bone; in old age (Fig. 5) there is bone resorption with little formation and consequently osteoporosis.

PROJECTION MICRORADIOGRAPHY

The development of this method is largely due to the excellent work of Cosslett & Nixon (Nixon & Cosslett, 1955; Nixon, 1957; Cosslett, Nixon & Pearson, 1957). This technique differs from the contact method

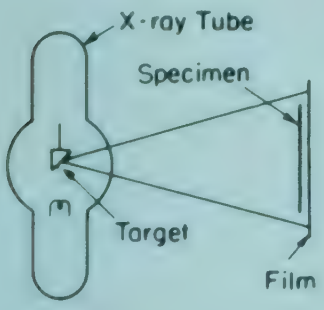


Figure 6. Diagram to show the position of the specimen in contact microradiography (after Cosslett).

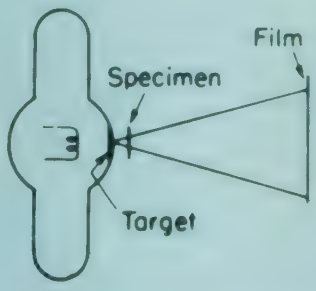


Figure 7. Diagram to show the position of the specimen in projection microradiography (after Cosslett).

(Fig. 6) in that the specimen lies in close relation to the X-ray source and at some distance from the photographic film (Fig. 7). A primary magnification is thereby obtained, and the image can then be magnified further photographically. In contradistinction to the contact method the range of resolution depends not on the fineness and even distribution of the photographic emulsion but on the sharpness of the initial image. This sharpness is achieved by having an exceedingly small focal spot, which in the machine in use in our department (kindly loaned by Mr Raymond Ely) is about 1 micron (Fig. 8). This is produced by the accurate focusing of an electron beam on a thin tungsten target by means of two electromagnetic lenses. The tungsten target which then directly emits the small X-ray beam forms the end-window of the X-ray tube. The limit of resolution is determined by the size of the focal spot, though if this is made too small the intensity of the X-ray beam is so reduced as to make

the exposure times impracticably long. So far the best resolution reported is 1,000 Å, or ten times better than the contact method.

As can be seen, the nearer the specimen lies to the source of X-rays the greater is the initial magnification. In practice good resolution has been obtained with the specimen about 1 mm. from the target (Figs. 10 and

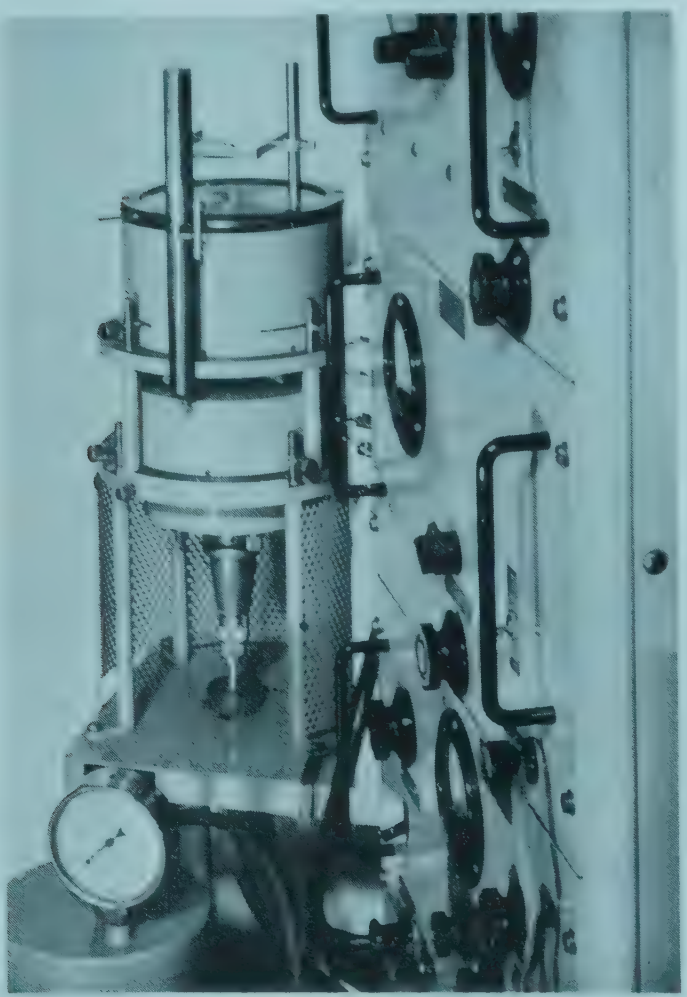


Figure 8. Machine used for projection microradiography. The bakelite top for the vacuum modification is shown in the lower left hand portion of the photograph, i.e. not in the position of use.

11). Biological specimens could not be placed closer because with lengthy exposures, which were sometimes necessary, the heat generated damaged the tissue or caused it to curl, with resultant loss of definition. Saunders

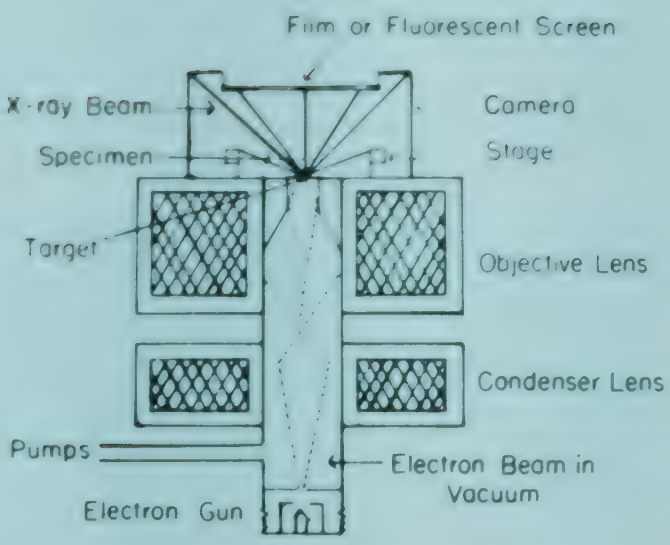


Figure 9. Diagram to show fine focusing of electron beam on target with production of X-rays from a source of about 1 micron (after Nixon).

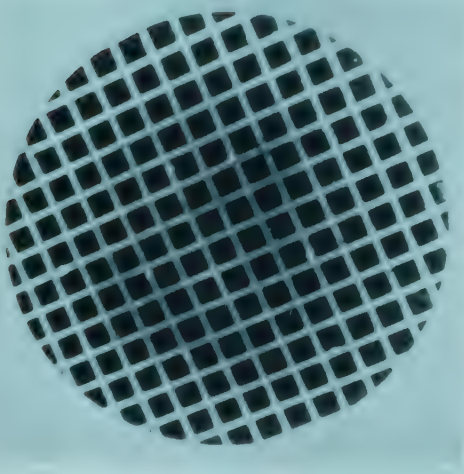


Figure 10. Showing resolution of 500-mesh grid, by projection microradiography.

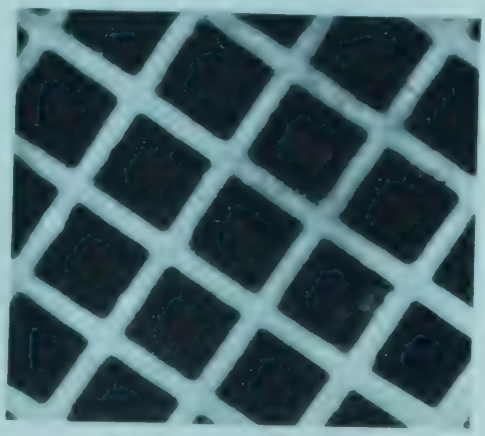


Figure 11. 2. Showing resolution of 1,500-mesh grid by projection microradiography.

(personal communication) recently modified a similar machine by introducing a vacuum system to enclose the specimen and plate (Fig. 12). This modification has been adopted in the instrument in use at the Royal College of Surgeons and has given much improved results. Freeze-dried sections of soft tissues about 10 μ in diameter are mounted on small brass rings (internal diameter 5 mm.) containing a thin film of Formvar

(Fig. 13). The film is made by dipping the ring into a 0·5 per cent solution of Formvar in ethylene dichloride and allowing it to dry. The ring containing the specimen is then carefully placed in the selected slot at the required distance from the target and the photographic plate laid on top of the brass table (Fig. 12). The whole is then covered with bakelite

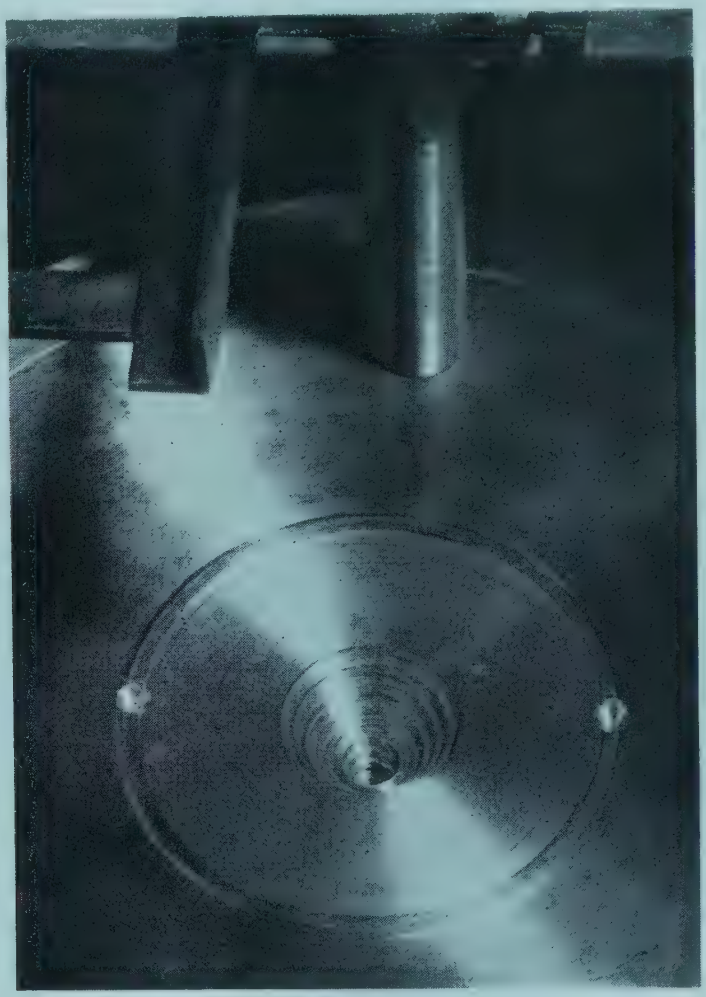


Figure 12. Photograph to show the modification necessary for a vacuum technique (Saunders). The target lies at the bottom of the series of slots and can just be seen in the centre. After the ring containing the specimen has been put in place, the photographic plate is laid on the top of the circular platform. The whole is covered with a bakelite top prior to evacuation.

top and sealed by a rubber washer prior to evacuation. By this technique the exposure time can be reduced, and most microradiographs are taken with X-rays generated at 7 kV. giving an exposure of 12 minutes.

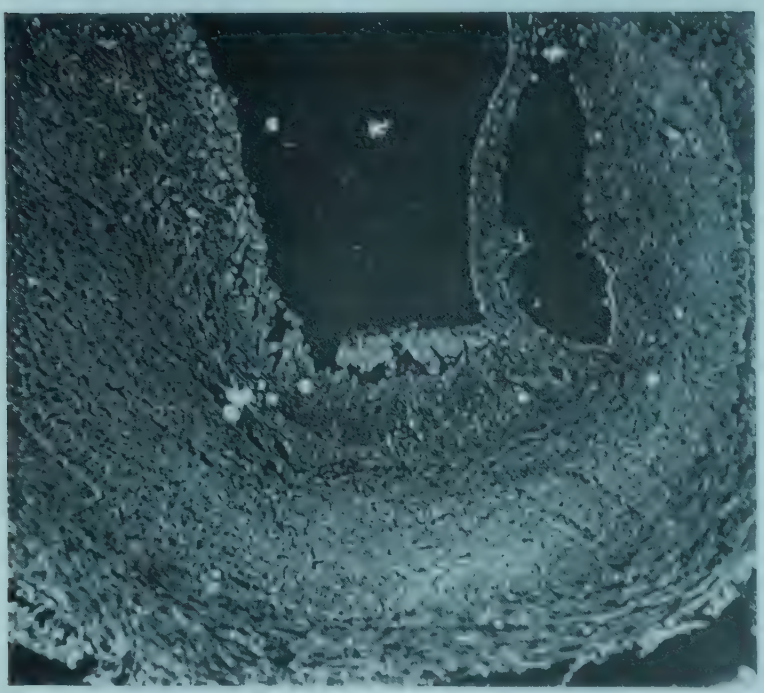
Microradiographs of arteries show the internal elastic lamina very clearly (Fig. 14, *a* and *b*); the arrangement of the collagen fibres at the edge of an old cardiac infarct is well displayed, and in some parts the cross-striations of the adjacent muscle fibres can be identified (Fig. 15, *a* and *b*). As anticipated, asbestos bodies in an area of scarred lung tissue are markedly opaque to X-rays (Fig. 16, *a* and *b*). A chance section through a portion of bronchial cartilage from an elderly person gave an excellent demonstration of the laying down of calcified particles (Fig. 17, *a* and *b*). The structure of bronchial epithelium can be well illustrated.

Figure 13. Brass ring of 5 mm. internal diameter used for mounting of histological freeze-dried section. Millimetre scale for comparison.

In addition to the cilia, a bar of phospholipid material not obvious in the photomicrograph is clearly outlined in the microradiograph (Fig. 18, *a* and *b*). Other minor changes can be detected in the epithelium and the method seems well-suited to supplement histological and histochemical studies on abnormalities in the bronchial mucosa. Microradiographs of other soft tissues so far taken are of interest and closely resemble photomicrographs in their main features.

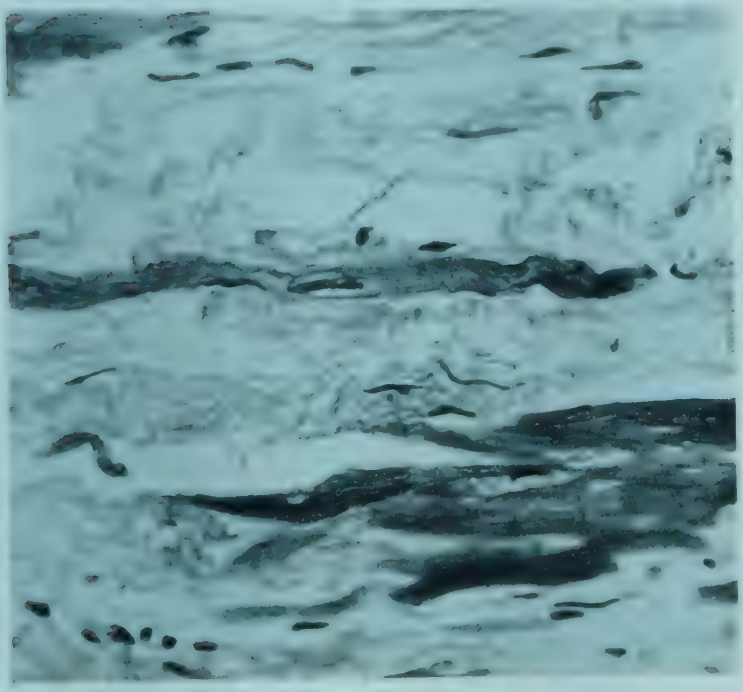
The point-focus projection X-ray microscope has many advantages.

(a)

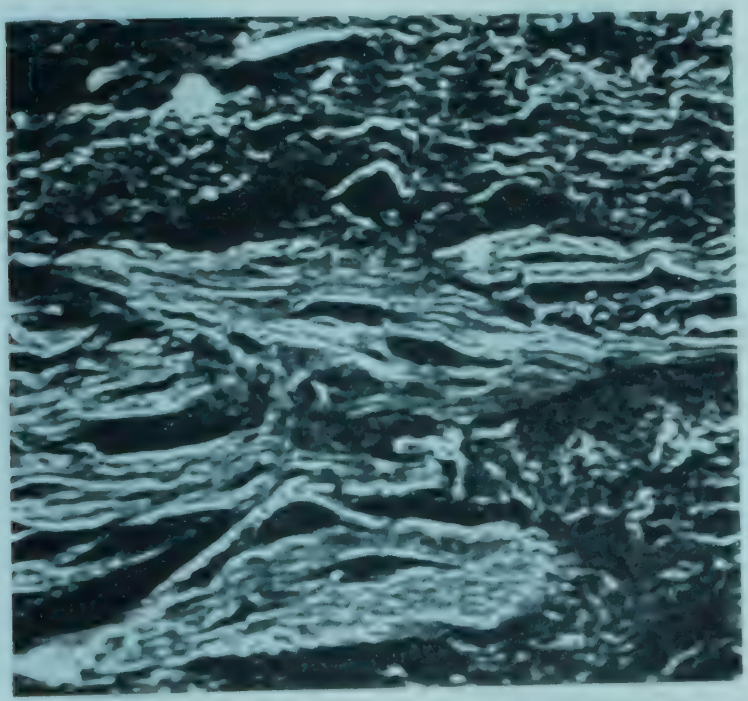


(b)

Figure 14. (*a*) × 100. Photomicrograph of portion of coronary artery with atheromatous thickening of the subintimal layer. (*b*) × 100. Microradiograph of same field shows in addition fragmentation of the internal elastic layer.



(a)



(b)

Figure 15. (a) × 400. Photomicrograph of edge of old cardiac infarct showing surviving heart muscle fibres and adjacent fibrous tissue. (*b*) × 400. Microradiograph of similar field shows arrangement of the collagen fibres, and cross-striation in the surviving muscle fibres

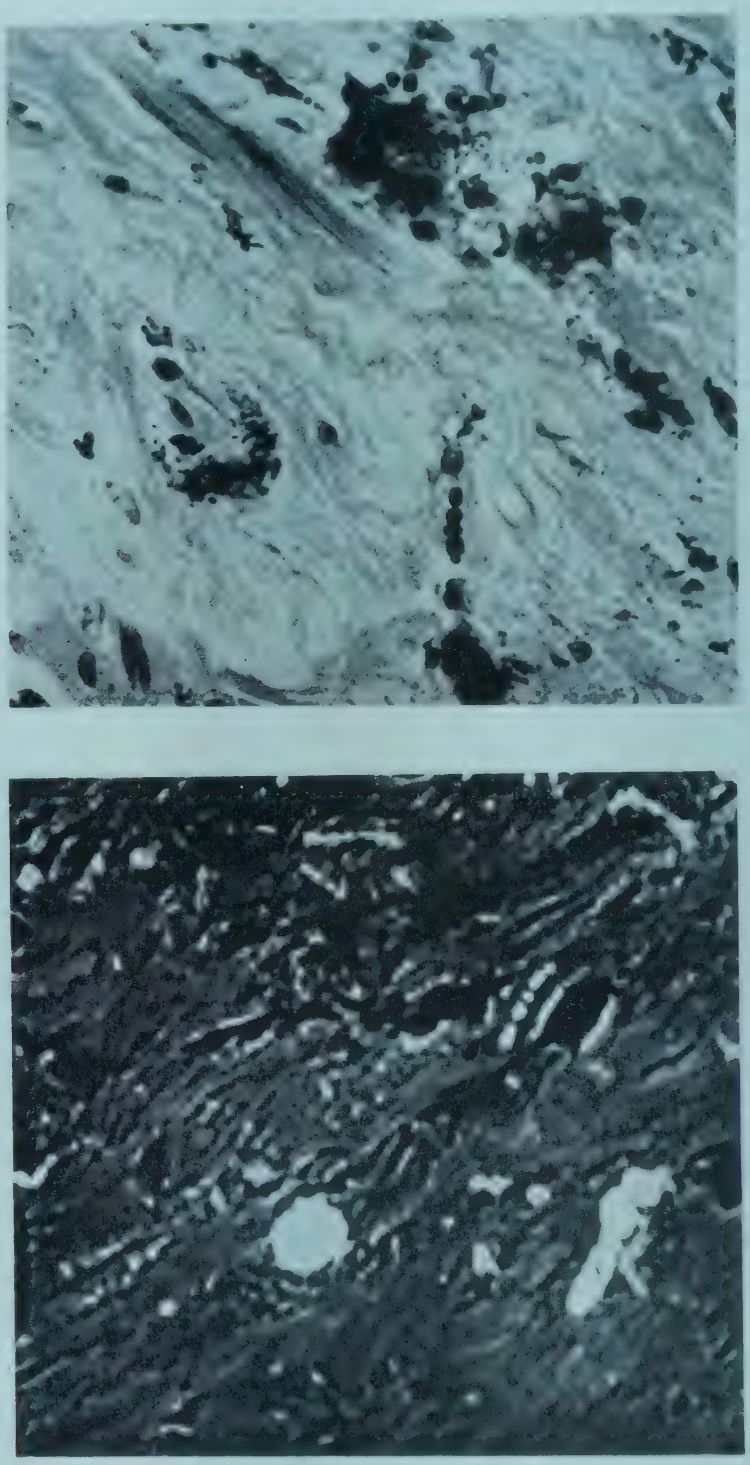
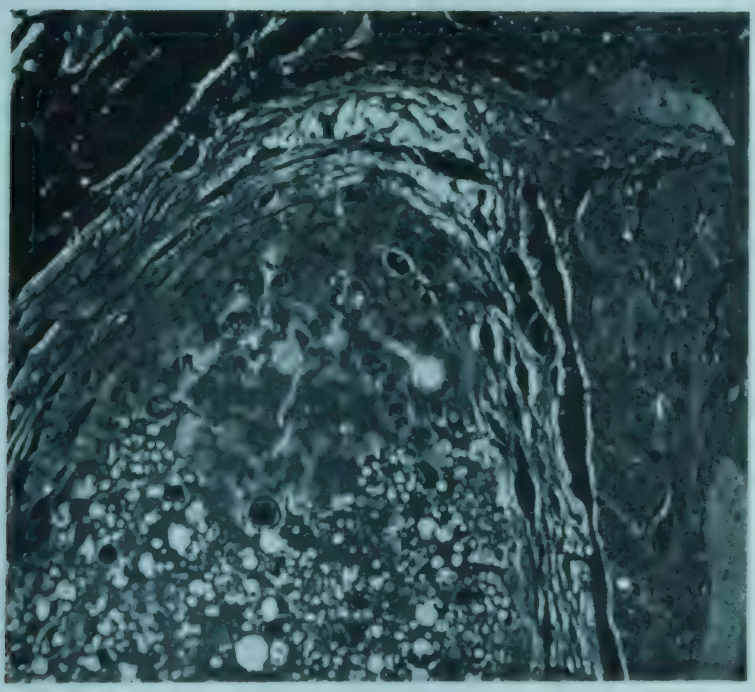


Figure 16. (*a*) × 500. Photomicrograph of scarred lung from a case of asbestosis. Small bodies can be seen. (*b*) × 500. Microradiograph of similar field shows small opaque segmented asbestos bodies.

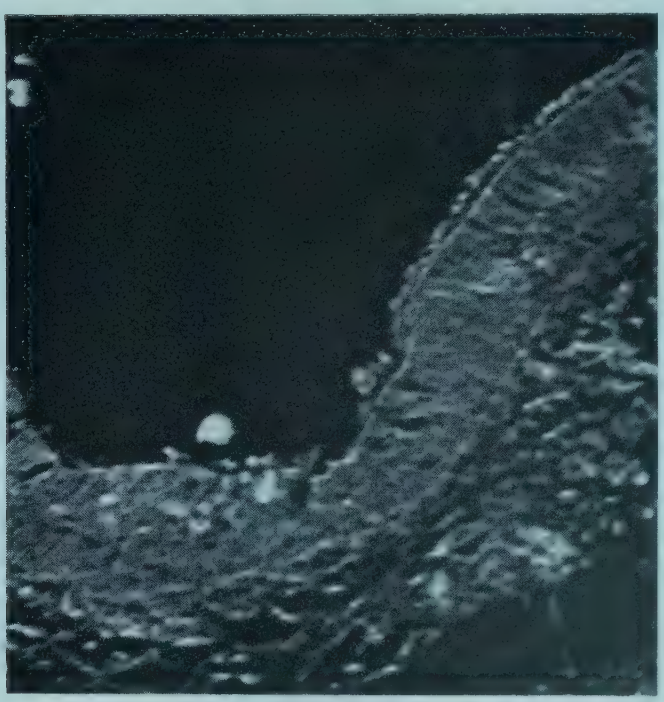
(a)



(b)

Figure 17. (*a*) × 120. Photomicrograph of bronchial cartilage undergoing calcification. (*b*) × 120. Microradiograph shows numerous small calcified particles being deposited as the cartilaginous matrix becomes replaced.

(a)



(b)

Figure 18. (*a*) ×400. Photomicrograph of ciliated epithelium from a main bronchus. (*b*) ×400. Microradiograph shows cilia and the characteristic phospholipid line at their bases.

In addition to providing a higher degree of resolution than the contact microscope it has greater flexibility. The area to be radiographed can be varied by altering the distance of the specimen from the target, and the tissue may be more easily positioned. As the image is produced by a point source of X-rays all depths of the section are in focus, and very thin sections are therefore not necessary. Stereographic pairs can thus be taken if a three-dimensional view is advantageous. Quantitative results can be obtained as with the contact method, whilst in addition the instrument can be adapted for microdiffraction. In biological specimens this method is limited to areas of tissue in which crystalline structures exist. Future developments are being explored with scanning techniques for the analysis of a specimen by differential X-ray absorption or by secondary emission.

REFLECTION X-RAY MICROSCOPE

Using the term in the strict sense this is the only true X-ray microscope in that it is the only one which employs a true optical system. By comparison with the two techniques previously described this method has made much less progress owing to technical difficulties which have yet to be overcome.

In spite of the ability of X-rays to penetrate matter it is well known that if they be made to strike a mirror surface at a very small angle they are totally reflected. Using an X-ray beam at an incident angle of about 1° it should theoretically be possible to obtain a maximum resolution of about 1 micron. In practice, however, owing to various optical aberrations it is necessary to use two or four mirrors in series in an attempt to correct these imperfections. Furthermore, in order to obtain such a resolution, mirrors with a very high degree of smoothness are necessary, and at present this poses great technical difficulties. Research on these matters is actively proceeding, though the method is not yet in general use for biological and medical problems. Much pioneer work has been done by Kirkpatrick in California (Kirkpatrick, 1957), who claims to have obtained a resolution of about 1 micron or one-tenth of the theoretical figure. More recently, Montel in France (1959) has incorporated prisms in a new reflection X-ray microscope.

CONCLUSION

It is likely that in future X-ray microscopy will become of the first importance in cellular research. Qualitatively, observations are likely to be useful in supplementing information already obtained with the light

microscope. The interpretation of electron micrographs has frequently given rise to difficulty partly because of the possibility of artifact production during the fixation and processing of the sections, and partly because the nature of objects observed at such high magnification has been in doubt. It is here that X-ray microscopy may be of the greatest value. Sections suitable for X-ray microscopy need no fixation and can be examined in the freeze dry state. The same section can be examined by the light microscope and comparison of individual fields made. When the range of resolution has been extended towards its theoretical maximum the microradiograph will be intermediate between the photomicrograph and the electronmicrograph, ideally placed for interpretation of the latter.

The projection method of X-ray microscopy is well suited for stereoscopic techniques which will provide a better visualization of tissue in depth. In spite of considerable technical difficulties, methods giving quantitative results are likely to be much more important. The amount and exact location of substances within a cell would both confirm and extend certain observations in histochemistry and cytochemistry.

Tissue structures poorly visualized in radiographs may be better shown by special histochemical staining methods prior to radiography. In the past, research in X-ray microscopy has suffered from the lack of suitable equipment available commercially. At present contact radiographic machines can be purchased, and in America a point-focus projection X-ray microscope has been produced by the General Electric Company in Milwaukee.

REFERENCES

Amprino, R. & Engström, A. (1952) Studies on X-ray absorption and diffraction of bone tissue. *Acta Anat.*, **15**, 1.

Barclay, A. E. (1947) Micro-arteriography. *Brit. J. Radiol.*, **20**, 394.

Barclay, A. E. (1949) Editorial article on A. E. Barclay. *Amer. J. Roentgenol.*, **62**, 119.

Barclay, A. E. (1951) *Micro-Arteriography and other Radiological Techniques employed in Biological Research.* Oxford.

Bellman, S. (1953) Microangiography. *Acta Radiol. Suppl.*, 102.

Bohatyrschuk, F. (1944) Über Ergebnisse der Mikroröntgenographie. *Acta Radiol., Stockholm*, **25**, 351.

Cheavin, W. H. S. (1947) Philatelic research by means of X-rays. *Year Book 29th Philatelic Congress*, pp. 72–80.

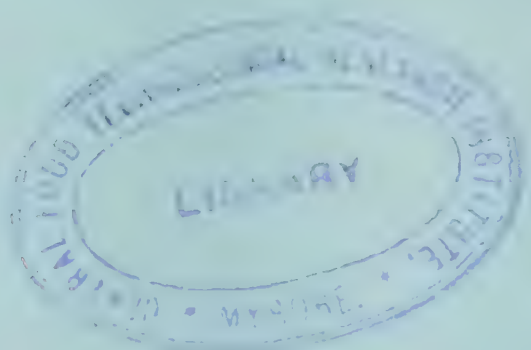
Cosslett, V. E. (1957) The scope of the symposium. *X-ray Microscopy and Microradiography*, Academic Press Inc., New York, 5.

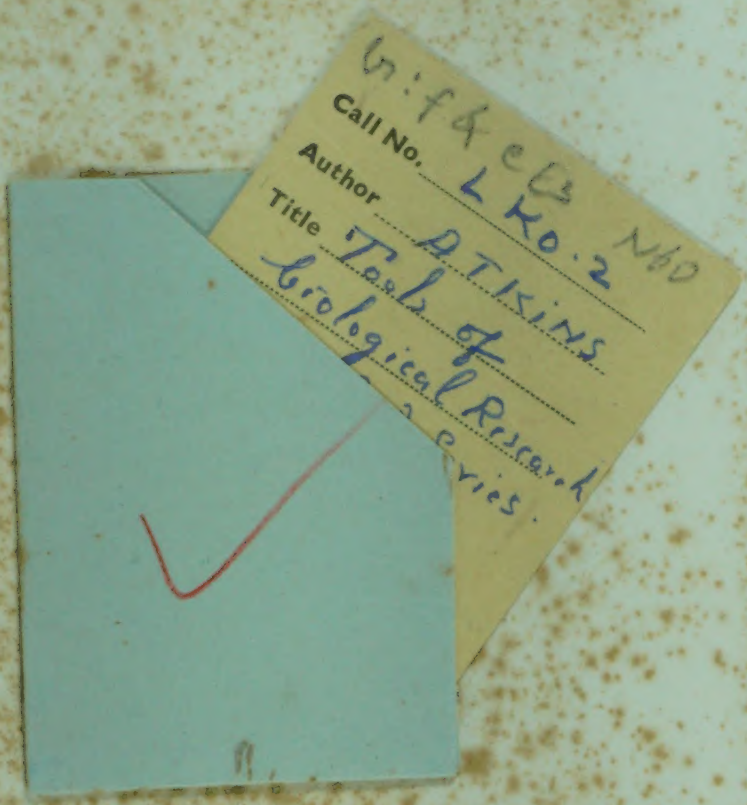
Cosslett, V. E., Nixon, W. C. & Pearson, H. E. (1957) Improvements in the point projection X-ray microscope. *X-ray Microscopy and Microradiography*. Academic Press Inc., New York, p. 96.

Cunningham, G. J. (1955) Studies in lung structure by low-voltage radiography. *Am. J. clin. Path.*, **25**, 253.

Cunningham, G. J. & Miller, J. W. (1952) A radiographic technique for the study of lung pathology. *Thorax*, **7**, 170.

Hauvillers, A. (1950) Réalisation de la microradiographie intégrale. *C.R. Acad. Sci., Paris*, **190**, 1287.
Eastman Kodak Company X-ray Division (1943) *Radiography of Materials*. Rochester, N.Y.
Engström, A., Bergman, S. & Bengt Engfeldt (1955) Microradiography. *Brit. J. Radiol.*, **28**, 517.
Engström, A. & Gothe, D. (1950) The mass of gastric mucosa cells measured by X-ray absorption. *Science*, **111**, 379.
Engström, A. & Lindström, B. (1950) A method for the determination of the mass of extremely small biological objects. *Biochim. et Biophys. Acta*, **4**, 351.
Fricke, H. (1932) Radiography of insects. *Radiogr. and clin. Photogr.*, **8**, 2.
Goby, P. (1913) Une application nouvelle des rayons X: la microradiographie. *C.R. Acad. Sci., Paris*, **156**, 686.
Kirkpatrick, P. (1957) The problem of reflection microscopy. *X-ray Microscopy and Microradiography*. Academic Press Inc., New York, p. 17.
Lamarque, P. (1938) Historadiography. *Brit. J. Radiol.*, **11**, 425.
Lefanu, W. R. (1959) Reynolds' two heads of John Hunter. *Brit. med. J.*, i, 580.
Lindström, B. (1955) Roentgen absorption spectrophotometry in quantitative cytochemistry. *Acta Radiol. Suppl.*, 125.
Montel, D. (1959) Nouvelles réalisations de microscope X à reflexion totale. Communication to *Second Int. Symp. on X-ray Microscopy and X-ray Microanalysis* (in the press).
Nixon, W. C. (1957) Point projection X-ray microscopy. *X-ray Microscopy and Microradiography*. Academic Press Inc., New York, p. 34.
Nixon, W. C. & Cosslett, V. E. (1955) Projection microradiography. *Brit. J. Radiol.*, **28**, 532.
Saunders, R. L. de C. H., Lawrence, J. & MacIver, D. V. (1957) Microradiographic studies of the vascular patterns in muscle and skin. *X-ray Microscopy and Microradiography*. Academic Press Inc., New York, p. 539.
Seeman, H. E. (1943) Miscellaneous applications of radiography and fluoroscopy. *Symposium on Radiography* 1936 *and* 1942, pp. 29–47.
Sherwood, H. F. (1934) The radiography of small biological specimens. *Radiogr. and clin. Photogr.*, **10**, 10.
Sherwood, H. F. (1936) The radiography of cloth. *J. Text. Inst. Manchester* (*Trans.*), **27**, T 162.
Sherwood, H. F. (1937) The radiography of small biological specimens. *J. biol. photogr. Ass.*, **5**, 138.
Sissons, H. A. (1950) Microradiography of bone. *Brit. J. Radiol.*, **23**, 2.
Sissons, H. A., Jowsey, Jenifer & Stewart, Louise (1959) The microradiographic appearance of bone tissue at various ages. *Communication Second Int. Symp. on X-ray Microscopy and X-ray Microanalysis* (in the press).
Sissons, H. A., Jowsey, Jenifer & Stewart, Louise (1959) Quantitative microradiography of bone tissue. *Communication Second Int. Symp. on X-ray Microscopy and X-ray Microanalysis* (in the press).
Tirman, W. S., Caylor, C. E., Barker, H. W. & Caylor, T. E. (1951) Microradiography. *Radiology*, **57**, 70.
Trueta, J., Barclay, A. E., Daniel, P. M. Franklin, K. J. & Prichard, M. M. L. (1947) *Studies in Renal Circulation*. Blackwell Scientific Publications, Oxford.
Wolff, A. L. (1950) Techniques of post-mortem angiography of the stomach. *Brit. J. Radiol.*, **23**, 8.




Call No. L KO. 2
Author ATKINS
Title Tools of
Research